

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:26 ; Search time 86.5 Seconds
(without alignments)
9941.972 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttattttatgatg.....aggcttttttctcctaataacc 2709

Scoring table:

BLAST62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09774490/runat.03082003.095415.25314/app.query.fasta_1.2887
-DB=A_Geneseq_19Jun03 -QWTF=fastan -SUFFIX=reg -MINMATCH=0.1 -LOOPFCU=0
-LIST=45 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human4.0.cdi
-MODE=LOCAL -OUTFMT=pct -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-USER=US09774490 @CEN 1.1.114 @runat.03082003.095415.25314 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FCAPOP=6
-FCAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4201	85.8	771	16	AAR71380	Human semaphorin I
2	4201	85.8	771	22	AAG62726	Amino acid sequenc
3	4201	85.8	771	23	ABG96413	Human ovarian canc
4	4194	85.7	796	19	ABY21264	Human semaphorin I
5	2562	52.3	477	16	AAR74175	Human collapsin.
6	2245.5	45.9	777	20	AAU27127	Human brain tissue
7	2245.5	45.9	777	21	AAU99427	Human PRO1491 (UNQ
8	2245.5	45.9	777	22	AAU29197	Human PRO polypept
9	2245.5	45.9	777	22	AAB66176	Protein of the inv
10	2245.5	45.9	777	24	ABU71285	Human PRO1491 prot
11	2245.5	45.9	777	24	ABU65742	Human secreted/tra
12	2245.5	45.9	777	24	ABU66075	Novel human secret
13	2245.5	45.9	777	24	ABU67579	Human secreted/tra
14	2245.5	45.9	777	24	ABU65437	Human PRO polypept
15	2245.5	45.9	777	24	ABU58573	Human PRO polypept
16	2245.5	45.9	777	24	ABU56109	Human secreted/tra
17	2245.5	45.9	777	24	ABU57104	Human PRO polypept
18	2245.5	45.9	777	24	ABU10683	Human secreted/tra
19	2209.5	45.1	807	22	AAG62729	Amino acid sequenc
20	2118	43.3	749	22	AAG62727	Amino acid sequenc
21	2044	41.7	775	24	ABR47588	Breast cancer asso
22	2030	41.5	775	19	AAW63748	Human semaphorin.
23	2024	41.3	775	20	AAU43090	Mouse semaphorin H
24	1942	39.7	777	20	AAU43091	Mouse semaphorin H
25	1942	39.7	777	22	AAG62731	Amino acid sequenc
26	1937	39.6	785	22	AAG62730	Amino acid sequenc
27	1849.5	37.8	751	21	AAW30617	Human semaphorin E
28	1849.5	37.8	751	21	AAW38379	Clone BR533.4. Ho
29	1849.5	37.8	751	23	ABP68623	Human pancreatic c
30	1847.5	37.5	751	22	AAG62728	Amino acid sequenc
31	1834	37.5	875	22	AAG65619	Novel human protei
32	1830	37.4	782	21	AAW23609	Human secreted pro
33	1830	37.4	782	21	AAG65620	Novel human protei
34	1826.5	37.3	785	21	AAW23636	Human secreted pro
35	1751.5	35.8	779	23	AAW84219	Amino acid sequenc
36	1751.5	35.8	779	23	AAW84219	Human ZSMF-16. Ho
37	1010	20.6	861	18	AAW17658	Mouse CD100 antige
38	1003	20.5	861	19	AAW58540	Human semaphorin.
39	1003	20.5	861	22	AAW10335	Murine CD100 amino
40	1003	20.5	861	22	AAW10351	Mouse CD100 protei
41	981.5	20.0	862	18	AAW17657	Human CD100 antige
42	981.5	20.0	862	22	AAW81036	Human CD100 amino
43	981.5	20.0	862	22	AAW51252	Human CD100 protei
44	966	19.7	834	22	AAW61238	Murine M-Sema-F pr
45	957.5	19.6	833	22	AAW03640	Human extracellular

ALIGNMENTS

RESULT 1
AAR71380
ID AAR71380 standard; Protein; 771 AA.
XX AAR71380;
AC AAR71380;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX Human semaphorin III protein.

XX Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
XX variola major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
XX Homo sapiens.
OS XX

PN W09507706-AL.
 XX
 PD 23-MAR-1995.
 XX
 PF 13-SEP-1994; 94WO-US10151.
 XX
 PR 13-SEP-1993; 93US-0121713.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 XX
 DR WPI; 1995-131177/17.
 DR N-PSDB; AAQ87442.
 XX
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 XX
 PS Example 2; Page 60-63; 101pp; English.
 XX
 CC The sequence of the human semaphorin III protein. The proteins
 CC encoded by the grasshopper semaphorin I (AAQ87441), human semaphorin
 CC III, vaccinia virus semaphorin IV (AAQ87443), Drosophila semaphorin I and
 CC II (AAQ87444-5), Tribolium semaphorin I (AAQ87446) or variola major
 CC (smallpox) virus semaphorin IV (AAQ87447) genes were used to generate a
 CC series of peptides (AA070370-R70418), which retain semaphorin receptor
 CC binding activity. The semaphorin derived or semaphorin receptor derived
 CC peptides are potent modulators of nerve cell growth, immune
 CC responsiveness and viral pathogenesis. They can be used in diagnosis and
 CC treatment of neurological disease and neuro-regeneration, immune
 CC modulation and diagnosis and treatment of viral and oncological infection
 CC and diseases.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 771 AA;

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR71380 (1-771)

QY 200 ATGGCTGGTTAACTAGGATTGTCTGCTTTTCTGGGAGTATTACTTACAGCAAGAGCA 259
 |||
 Db 1 MetGlyTyrLeuThrArgileValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 |||
 QY 260 AACTATCAGAAATGGGAGAACAAATGTGCCAGCGTGAATATCTTACAAAGAAATGTTG 319
 |||
 Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLeuLeuSerTyrLysGluMetLeu 40
 |||
 QY 320 GAATCCAAACAAATGATCACTTCAATGCTTGGCCCAACAGCTCCAGTATCATCACTTC 379
 |||
 Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 |||
 QY 380 CTTTGTGGATGAGCAACGGAGTAGGCTGTATGTTGGAGCAAGGATGCACATATTTTCATTC 439
 |||
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 |||
 QY 440 GACCTGGTTAAATATCAGGATTTTCAAAGATTGTGGCCAGTATCTTACACCAAGAAGA 499
 |||
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 |||
 QY 500 GATGAATCAAGTGGCTCGAAAAGACATCTCGAAGAGATGCTAAATTTTCATCAAGGTA 559
 |||
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelleyVal 120
 |||
 QY 560 CTTAAGGCATATAATCAGACTCCTGCTGTACGCTGTGGAAACGGGGCTTTTCATCCAATT 619
 |||

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProfile 140
 |||
 QY 620 TGCACCTACATTGAAATTGGACATCATCTCAGGACATATATTTTAACTCGAGACTCA 679
 |||
 Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 |||
 QY 680 CATTTTGAACCGCGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGCAGCATCCCTT 739
 |||
 Db 161 HisPheGluAsnGlyArgGlySerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 |||
 QY 740 TTAATAGATGAGAAATTATATCTCTGGAACTCGAGCTGATTTTATGGCGCAGACTTTGCT 799
 |||
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 |||
 QY 800 ATCTTCCGAACCTCTTGGCACACCCATCAGGACAGGACAGCATGATTCAGAGTGG 859
 |||
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 |||
 QY 860 CTCATATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCCTCAAGAT 919
 |||
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 |||
 QY 920 GACAAAGTATATCTTTTCTCCGTGAAATCAATAGATGAGAACACTCTCGAAAGACT 979
 |||
 Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 |||
 QY 980 ACTACGCTAGAAATAGTGCAGATATGCAGATATGACATTTTGGAGGCGACAGAGTCTGGTG 1039
 |||
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 |||
 QY 1040 AATAAATGGACAACTCTCTCAAGCTCGTCTGATTTTCTCAGTCCAGGTCCTCAATGGC 1099
 |||
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 |||
 QY 1100 ATTCACTCATTTTGTGAACTGCAGATGTATCTCTTAATGAATTTTAAAGATCTCTAAA 1159
 |||
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 |||
 QY 1160 AATCCAGTTGTATATGGAGTGTATTCAGACTTCCAGTAACTATTTCAAGGGATCAGCCGTG 1219
 |||
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 |||
 QY 1220 TGTATGTATAGCATGAGTGTGAGAGGGGTGTTTCTTGTGCTCATATGCCACAGGGAT 1279
 |||
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 |||
 QY 1280 GGACCCCAACTATCAATGGTGCCTTATCAAGAGAGTCCCTATFCCACGCCAGGAACT 1339
 |||
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 |||
 QY 1340 TGTCCACGACAAACATTTGGTGGTTTGTACTCTCAAGGACCTTCTCTGATGATGTTATA 1399
 |||
 Db 381 CysProSerTyrThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 |||
 QY 1400 ACCTTTTGAAGAAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAAACAATCGCCA 1459
 |||
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 |||
 QY 1460 ATAGTGATCAAAACCGATGTAATATCAATTTACAAATTTGCTGTAGACCGAGTGTGAT 1519
 |||
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 |||
 QY 1520 GCAGAGATGACAGATGATGTTATGTTTATCGACACAGATGTGGGACCGTCTCTTAAA 1579
 |||
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 |||
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCCTCTCGAAGAAATG 1639
 |||
 Db 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
 |||
 QY 1640 ACAGTTTTCGGGAACCGACTGCTGTTTTCAGCAATGAGAGCTTTTCCATTAACGACGACAA 1699
 |||
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 |||

PN	XX	WO200138491-A2.
PP	XX	
PD	XX	31-MAY-2001.
PX	XX	
PF	XX	07-NOV-2000; 2000WO-US41943.
PP	XX	
PR	XX	08-NOV-1999; 99US-0164056.
XX	XX	(GEHO) GEN HOSPITAL CORP.
PA	XX	
PI	XX	Behar O, Woolf CJ;
XX	XX	
DR	XX	WPI; 2001-451494/48.
PT	XX	
PT	XX	Polypeptide sequences that encompass the hanatoxin-like sequences of
PT	XX	semaphorins, useful as a drugs to treat any condition or disease that
PT	XX	is characterized by abnormal calcium channel function -
XX	XX	
PS	XX	Claim 6; Page 11; 29pp; English.
XX	XX	
CC	XX	The present sequence represents a semaphorin. The specification
CC	XX	describes hanatoxin-like sequences (HTLS) found in the semaphorin
CC	XX	domain of mammalian secreted semaphorins. Hanatoxin is a tarantula
CC	XX	toxin that selectively blocks some voltage-gated potassium and calcium
CC	XX	channels. The HTLS is responsible for the dorsal root ganglion repulsion
CC	XX	and growth cone collapse activities associated with semaphorins.
CC	XX	Polyptides containing HTLS can be used to modulate the activity of
CC	XX	calcium channels. The peptides can also be used as an antigen to
CC	XX	generate antibodies that can then be used to modulate the activity of
CC	XX	calcium channels by inactivating naturally occurring channel ligands.
CC	XX	The peptides or antibodies can be used as drugs to treat any condition
CC	XX	or disease that is characterized by abnormal calcium channel function.
XX	XX	
SQ	Sequence	771 AA;
Alignment Scores:		
Pred. No.:	0	Length: 771
Score:	4201.00	Matches: 771
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	85.80%	Indels: 0
DB:	22	Gaps: 0
US-09-774-490-1 (1-2709) x AAG62726 (1-771)		
Qy	200	ATGGCGTGTAACTAGGATTGTCGTCTTTTCTGGGAGATTACTTACAGAAGACA 259
Db	1	MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaAargAla 20
Qy	260	AACNTACGAATGGGAGAACATGTGCCAAGCGTGAAATATCTCTACAAGAAATGTG 319
Db	21	AsnTyrglnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyLysGluMetLeu 40
Qy	320	GAATCCAACAATGTGATCACTTTCAAATGGCTTGGCCAACAGCTCCAGTTATCATACCTTC 379
Db	41	GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSeryHisThrPhe 60
Qy	380	CTTTTGGATCAGGAACGGATGAGCTGTATGTTCGAGCAAAAGGATCACATATTTTCATTC 439
Db	61	LeuLeuAspGluAuArgSerArgLeuTyValGlyAlaLysAspHisIlePheSerPhe 80
Qy	440	GACCTGTTTAATACRAGGATTTTCAAAAGATTGTGTGGCCASTATCTTACACCAGAAGA 499
Db	81	AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyThrArgArg 100
Qy	500	GATCAATGCAGTGGGCTGGAAGAAGACATCTCGAAAGAATGTGCTAATTTTCATCAAGGTA 559
Db	101	AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
Qy	560	CTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGGAACCGGGGCTTTTCATCCAATY 619
Db	121	LeuLysAlaTyArgAsnGlnThrHisLeuTyArgAlaCysGlyThrGlyAlaPheHisProIle 140

QY	380	CTTTTGGATGAGAAACGGAGTAGCTGTATGCTGGAGCAAAAGGATCACATATTTTCATTC	4339
Db	61	LeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe	80
QY	440	GACCTGGTTAAATATCAAGGATTTTCAAAGATGCTGGCCAGTATCTTACACCAAGAAGA	499
Db	81	AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg	100
QY	500	GATCAATGCAAGTCGGCTGCAAAAGACATCTCGAAAGAATGTGCTAAATTTTCATCAAGGTA	559
Db	101	AspGluCysLysIleTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal	120
QY	560	CTTAAGGCATATAATCAGACTCACTGTGTAGCCCTGTGGAACGGGGGCTTTTCATCCAAVT	619
Db	121	LeuLysAlaTyrAsnGlnThrHisLeuTyraLacysGlyThrGlyAlaPheHisProIle	140

620 TGCACCTACATTGAATGGACATCATCTCGAGGACAAATATTTTAAGCTGGAGAACTCA 679
141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLeuLeuGluAsnSer 160
680 CATTTTGAAGACCGCGTGGAGAGTCCATATGACCCCTAAGCTCTCTCACAGCATCCCTT 739
161 HisPheGluAsnGlyArgGlyLeuSerProTyrAspProLeuLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTCTGAACTGACGCTGATTTTATGGCGGAGACTTTGCT 799
181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAGACTTGGGACACACCAACCAATCAGACAGACAGCATGATTCCAGGTGG 859
201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
860 CTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCCTCAAGAT 919
221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTCTTCGTAAGTCAATAGATGAGAGACACTCTGGAAGCT 979
241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGAAATAGGTGAGATATGCAAGATGACTTTGGAGGCGACAGAACTCTGGTG 1039
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAAATGGACAACTTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCAAATGCG 1099
281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTGAAGTGCAGGATGATTCCTTAATCAACTTTAAGATCCTAAA 1159
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 AATCAGTTGTATATGGAGTGTTCAGCTCCAGTAACTTTCAAGGATCACCGCTG 1219
321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGTATAGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCACAGGAT 1279
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GACCCCACTATCAATGGTGCCTTATCAAGAGAGTCCCTATCCAGCGGCAGAACT 1339
361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCACGACAAACATTTGGTGGTTTGTACTCTCAAGAGACCTTCTCGATGATGTATA 1399
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAAGAGTCAATCCAGTGTATCAATCCAGTGTTCCTATGAACAATCCGCCA 1459
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
1460 ATAGTATCAAAACCGATGTAATTAATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGCTCTTAAA 1579
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGTTCCTGTGGAAGAAATG 1639
461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluMet 480
1640 ACAGTTTTCGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATTGGTTCAACGCTGGGTTGGCCAGTCCCTTTTACACCGGTGTGATATTTAC 1759

501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAACGCTGTCTGAGTGTTCCTGCCCGACACCTTACTGCTGTGGATGGTTCCT 1819
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAACAGCACCAAGATATAAGAAAT 1879
541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
1880 GGAGACCCACTGACTCTGTTTACAGCTTACCATGATATCAATCATCCATGGCCACAGCCT 1939
561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
1940 GAAGAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGGATGAGTCCCGAAG 1999
581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGCAATGAAGACGCAAAAGAGAG 2059
601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgArgGlnGluArgLysGluGlu 620
2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTACA 2119
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
2120 CAGAAGATTCCAGCAATTTACTCTGCTCATCGGTGGAACATGGGTTCATACAACTCTT 2179
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
2180 CTTAAGGTAAACCCGGAAGTCATTGACACAGAGATTTGGAAGAACTTCTTCATAAAGAT 2239
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
2240 GATGATGGAGATGCTCTTAAGACCAAAAGAAATGTCCAATAGCATGCACACCTAGCCAGAAG 2299
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
2300 GTCTGGTACAGAGACTTCTACGCTCATCAACACCCCAATCTCAACACCATGATGAG 2359
701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
2360 TTCTGTGAACAAGTTTGGAAAAGGACCGAAAACAAACGTCGGCAAGCCAGGACATACC 2419
721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
2420 CCAGGGAACAGTAAACAATGGAAGCCTTCAAGAAAATAAGAAAGGTAGAAAACAGGAGG 2479
741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
2480 ACCCAGCAATTTGAGGGGACCCAGGAGTGTCT 2512
761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 3
ABG96413
ID ABG96413 standard; Protein; 771 AA.
XX
AC ABG96413;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M473.
XX
Human; ovarian cancer; marker; cancer; familial history; brain disorder;
central nervous system disorder; bacterial meningitis; viral meningitis;
Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
brain herniation; inflammation; encephalitis; testicular disorder;
nontuberculous granulomatous orchitis; connective tissue disorder;
heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
histological type; carcinogenic; ovarian cancer marker.

OS Homo sapiens.
 PN WO200271928-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 14-MAR-2002; 2002WO-US07826.
 XX
 PR 14-MAR-2001; 2001US-276025P.
 PR 14-MAR-2001; 2001US-276026P.
 PR 10-AUG-2001; 2001US-311732P.
 PR 19-SEP-2001; 2001US-323580P.
 PR 26-SEP-2001; 2001US-324967P.
 PR 26-SEP-2001; 2001US-325102P.
 PR 26-SEP-2001; 2001US-325149P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
 PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX
 DR WPI; 2002-723277/78.
 DR N-PSDB; ABS76512.
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient -
 XX
 PS Disclosure; Page 401-402; 481pp; English.
 XX
 CC The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. bacterial or viral meningitis or encephalitis),
 CC connective tissue disorders (e.g. nontuberculous granulomatous orchitis),
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer,
 CC determining whether ovarian cancer has metastasized or is likely to
 CC metastasize, selecting a composition for inhibiting ovarian cancer,
 CC assessing the ovarian carcinogenic potential of a compound, or
 CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
 CC present amino acid sequence represents one of the ovarian cancer markers
 CC described in the invention.
 XX
 SQ Sequence 771 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 23 Gaps: 0
 US-09-774-490-1 (1-2709) x ABG96413 (1-771)
 QY 200 ATGGGCTGGTAACTAGAGATTGCTCTTTCTGGGGAGTATTACTTACAGCAGAGCA 259
 |||||||

Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGGTGAAATTTATCTCTACAAAGAAATGTTG 319
 |||||||
 Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCAAATGTGATCATCTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
 |||||||
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGTGATGAGGAACCGAGTAGGCTGTATCTTGGAGCAAGAGATCATATTTTCATTC 439
 |||||||
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisPheSerPhe 80
 QY 440 CACCTGGTTAATCAAGGATTTTCAAGAGATTGTGGCCAGTATCTTACACCAAGAGA 499
 |||||||
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTTAATTTTCATCAAGTA 559
 |||||||
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATATCATCACTCCTGTACGCTGTGGAAACGGGGCTTTTCATCCATTT 619
 |||||||
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACTACATGTAATTTGACATCATCTCTGAGACAAATATTTTAACTGGAGAACTCA 679
 |||||||
 Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAACACGGCCGTGGGAAGAGTCCATATCACCTTAAGCTGCTGACAGCATCCCTT 739
 |||||||
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTCCT 799
 |||||||
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCGGAACCTCTGGGCGACACCCACCAATCAGGACAGACAGCATGATTCAGGTGG 859
 |||||||
 Db 201 IlePheArgThrLeuGlyHisHisHisPheIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCCAAAGTTCATTTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
 |||||||
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACTCTCTGGAAGAGCT 979
 |||||||
 Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCGTAGAATAGATGATGCAAGATGCAATGACCTTTGGAGGCGCAGAGTCTGGTG 1039
 |||||||
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGAACAATCTTCAAGAGCTGTCTGATTTGCTCAGTCCAGGTGCCAAATGGC 1099
 |||||||
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTGTAAGTGCAGGATGATCTCTTAATGACTTTAAAGATCTCTAAA 1159
 |||||||
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGTATGAGTGTGTTTACGACTTCCAGTAACATTTTCAAGGAGATCAGCGGTG 1219
 |||||||
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGATGTATAGCATGAGTGTGTGAGAAAGGTGTCTCTGGTCCATATGCCACAGGAT 1279
 |||||||
 Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisAsp 360
 QY 1280 GGACCCAACTATCATGGGTGCTTATCAAGGAAGAGTCCCTTATCCACGCCAGGAACT 1339
 |||||||
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

1340 TGTCCAGCAAAACATTTGGTGGTTTGAATCTTACAAAGGACCTTCCCTGATGATCTTATA 1399
Db
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAATCCGCCA 1459
Db
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGGATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGAGTCAGATCATCTGTTATCTTATCGGAACAGATGTTGGGACCGTCTCTAAA 1579
Db
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db
461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuGluGluMet 480
QY 1640 ACAGTTTTCGGAACCGACTCTATTTCCAGCAATGGAGCTTTCACCTAAGCAGCAACA 1699
Db
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGCTCGGTTGCCAGCTCCCTTTACACGGTGTGATATTAC 1759
Db
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTGGATGTTCT 1819
Db
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTCCCTACTGCAAGAGACGACAGCAGCAGATATAAGAAAT 1879
Db
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGTTTACAGCTTACACCATGATTAATCACCAGCCGACAGCCCT 1939
Db
561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGTTAGAGATAGTAGACATTTTGGATGTCAGTCCGAAG 1999
Db
581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCCCTGCTTATTTGGCAATTCAGAGCGCAATTAAGAGCGCAAAAGAGAG 2059
Db
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTCGATCATATCATCATCAGGACAGATCAAGCGCTTCTGCTAGTGTCTACAA 2119
Db
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATACCTCTGCCATGCGGTGGAACTGGTTCATACAACTCTT 2179
Db
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAGGTAAACCTGGAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTCATAAAGAT 2239
Db
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGTGGCTTCTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGAG 2299
Db
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTCAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGATGGATGAG 2359
Db
701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGACAGATTGGAAAGGACCGAAACACACGTCGCGCAAGCCAGGACATACC 2419
Db
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

2420 CCAGGGAACAGTAACAAATGGAAGCAGCTTCAAGAAATAGAAAGGTAGAAACAGGAGG 2479
Db
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACAGAAATTTAGAGGGGACCCAGGAGTGTC 2512
Db
761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
ID AAY21264 standard; Protein; 796 AA.
XX AC AAY21264;
XX DT 22-JUL-1999 (first entry)
XX DE Human semaphorin III wild type protein fragment 1.
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
frame-shift mutation; age-related disease; neurodegenerative disorder;
Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
Huntington's disease; multiple sclerosis; alcoholic liver disease;
diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
high mobility group protein-C; neuroendocrine specific protein A.
XX OS Homo sapiens.
XX PN WO9845322-A2.
XX PD 15-OCT-1998.
XX PF 02-APR-1998; 98WO-IB00705.
XX PR 10-APR-1997; 97US-0043163.
XX PA (UYUT-) RIJKSUNIV UTRECHT.
XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI; 1998-609901/51.
XX DR N-PSDB; AAX75767.
XX PT Diagnosing disease by detecting frameshift mutations in RNA or
corresponding protein mutations - used to diagnose cancer and
neurological diseases, particularly Alzheimer's disease, and also
for treatment and prevention with specific ribozymes or wild-type
RNA
XX Disclosure; Figure 16; 258pp; English.
XX This invention describes a novel method for the diagnosis of a disease
caused by, or associated with, an RNA molecule that has a frameshift
mutation. The method is used to diagnose age-related diseases, especially
cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
and many others listed) or susceptibility to these disorders. The method
allows a definitive diagnosis of Alzheimer's disease in living patients,
at an early stage. It is based on the observation that disease may be
caused by mutations in RNA rather than DNA. The invention describes the
use of neuronal system RNA molecules, specifically proteins including
beta-amyloid precursor protein (beta-APP), the microtubule associated
proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group

CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 796 AA;

Alignment Scores:

Pred. No.: 0 Length: 796
Score: 4194.00 Matches: 773
Percent Similarity: 97.48% Conservative: 0
Best Local Similarity: 97.48% Mismatches: 0
Query Match: 85.66% Indels: 20
DB: 19 Gaps: 1

US-09-774-490-1 (1-2709) x AAY21264 (1-796)

QY 194 TCGACATGGCTGGTAACTAGGATGCTGCTCTTTCTGGGGAGTATTACTACAGCA 253
DB 4 CysSerMetGlyTTrpLeuThrArgileValCysLeuPheThrGlyValLeuLeuThra 23
QY 254 AGAGCAAACTATCAGAAATGGGAGAACAAATGTGCCAAGCTGMAATTTATCTACAAAGAA 313
DB 24 ArgAlaAsnTyrGlnAsnGlyLysAsnValProArgLeuLeuLysSerTyrLysGlu 43
QY 314 ATGTTGGAAATCAACAATGTGATCACTTTCAATGCTTGGCCCAAGCTCCAGCTTATCAT 373
DB 44 MetLeuGluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHis 63
QY 374 ACCTTCCTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 433
DB 64 ThrPheLeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePhe 83
QY 434 TCATTGACCTGGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTACACC 493
DB 84 SerPheAspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThr 103
QY 494 AGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAGATGCTAAATTTTCATC 553
DB 104 ArgArgAspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelle 123
QY 554 AAGTACTTAAGGCATATAATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTCAT 613
DB 124 LysValLeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHis 143
QY 614 CCAATTCGACCTACATTTGAATTTGGACATCATCTCTGAGGCAATATTTTAAAGCTGGAG 673
DB 144 ProIleCysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGlu 163
QY 674 AACTCACATTTGAAACGGCGTGGGAGAGTCCATATGACCTTAAAGCTCTGACAGCA 733
DB 164 AsnSerHisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAla 183
QY 734 TCCCTTTTAAATAGATGAGAAATTAATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGAC 793
DB 184 SerLeuLeuIleAspGlyLeuLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAsp 203
QY 794 TTGCTATCTTCGGAACCTCTGGGACCAACCCCAATCAGACAGACGATGATTC 853
DB 204 PheAlaIlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSer 223
QY 854 AGTGGCTCAATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGCACATCCT 913
DB 224 ArgTrpLeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnPro 243
QY 914 GAAGATGCAAGATATATCTTTCTTCGTCGAAATGCAATAGATGGAGAACACTCTGGA 973
DB 244 GluAspAspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGly 263
QY 974 AAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGATTTTGGAGGCCACAGAGT 1033
DB 264 LysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSer 283
QY 1034 CTGGTGAATAATGACACATCTCTCAAGCTCGTCTGATTTGCTAGTCCAGGTCCA 1093
DB 284 LeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyPro 303

QY 1094 AATGGCATTGACACTCATTTTGTATGAAGTGCAGGATGTATTCTTAATGAACCTTTAAAGAT 1153
DB 304 AsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAsp 323
QY 1154 CCTAAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTAACATTTTCAAGGGATCA 1213
DB 324 ProLysAsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySer 343
QY 1214 GCCGTGTGTATAGCATGATGATGTGAAGAGGGTGTTCCTTGGTCCATATGCCCCAC 1273
DB 344 AlaValCysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHis 363
QY 1274 AGGATGAGCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCCTATCCACGGCCA 1333
DB 364 ArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgPro 383
QY 1334 GGAATCTGTCCACAAAACATTTGGTGGTTCCTGACTCTACAAAGGACCTTCTGTGATCAT 1393
DB 384 GlyThrCysProSerLysThrPheGlyLysPheAspSerThrLysAspLeuProAspAsp 403
QY 1394 GTTATAACCTTTGCAAGAAGTCCATCCAGCCATGTACAATCCAGTGTTCCT- 1444
DB 404 ValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsn 423
QY 1445 -----ATGAACAAT 1453
DB 424 ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValMetAsnAsn 443
QY 1454 CGCCCAATAGTGTACAAAACGGATGTAATATCAATTTACACAAATTTGCTAGACCGA 1513
DB 444 ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArg 463
QY 1514 GTGATGTCAGAAAGATGACAGTATGATGTTATCGGAACAGATGTTGGGACCGTT 1573
DB 464 ValAspAlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrVal 483
QY 1574 CTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTGCTGAA 1633
DB 484 LeuLysValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGlu 503
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGCTATTTTCCAGCAATGGAGCTTCCACATAGCAG 1693
DB 504 GluMetThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGln 523
QY 1694 CAACAATATATATTGTTCAACGGCTGGGTCCAGCTCCCTTACACCGGTGTGAT 1753
DB 524 GlnGlnLeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAsp 543
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTTCCTCCCGAGACCCCTTACTGTGCTGGGAT 1813
DB 544 IleTyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAsp 563
QY 1814 GGTTCGTGATGTTCTCGCTATTTTCCACTGCAAGAGACGCCACAGACGACAGATATA 1873
DB 564 GlySerAlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIle 583
QY 1874 AGAATGAGACCCACCTGACTCTGCTGACATTCACCATGATATACCATGGCCAC 1933
DB 584 ArgAsnGlyAspProLeuThrHisCysSerAspLeuHisIleAspAsnHisIleGlyHis 603
QY 1934 AGCCTCGAAGAGAAATCATCTATGGTGTAGATAATAGTAGACATTTTTCGAATGCACT 1993
DB 604 SerProGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSer 623
QY 1994 CCGAAGTCCGACAGAGCGCTGCTATTTGGCAATTCAGAGCGCAATTAAGAGCGCAAAA 2053
DB 624 ProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLys 643
QY 2054 GAAGAGATCAGCTGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGGTAGT 2113
DB 644 GluGluIleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSer 663

QY 2114 CTACACAGAGGATTTCAGGCAATACCTCTCCATGGGTGGAACATGGGTTTACAAA 2173
 DB 664 LeuGlnGlnLysAspSerGlyAsnTrpLeuGlnLysHisAlaValGluHisGlyPheIleGln 683
 QY 2174 ACTCTCTTAAAGTAACCTCGAAGTCATTTGACACAGAGCAATTTGGAAGAATCTTCTCAT 2233
 DB 684 ThrLeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHis 703
 QY 2234 AAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCATAGATGACACCTAGC 2293
 DB 704 LysAspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSer 723
 QY 2294 CAGAAGGCTCTGATACAGAGCTTCATGACGCTCATCAACCCCACTCAACACGATG 2353
 DB 724 GlnLysValTrpTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMet 743
 QY 2354 GATGAGTCTGTGAACAAGTTTGGAAAAGGGACCGAAAACAACGTCGGCAAGGCCAGGA 2413
 DB 744 AspGluPheCysGluGlnValTrpLysArgAspArgLysGlnArgProGly 763
 QY 2414 CATACCCAGGCAACAGTACAAATGGAAGCACTTACAAGAAATAAGAAAGGTAGAAC 2473
 DB 764 HisThrProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsn 783
 QY 2474 AGGAGGACCCAGCAATTTGAGAGGGCCACCCAGGAGTGTC 2512
 DB 784 ArgArgThrHisGluPheGluArgAlaProArgSerVal 796

RESULT 5

AAR74175
 ID AAR74175 standard; Protein; 477 AA.

XX AC AAR74175;

XX DT 01-NOV-1995 (first entry)

XX DE Human collapsin.

XX KW Collapsin; antibody; therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Binding-site 9..19 /note= "antibody binding site"

FT Binding-site 51..65 /note= "antibody binding site"

FT US5416197-A.

XX PD 16-MAY-1995.

XX PF 15-OCT-1993; 93US-0136922.

XX PR 15-OCT-1993; 93US-0136922.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Luo Y, Raper JA;

XX DR WPI; 1995-193478/25.

XX DR N-PSDB; AAQ92331.

XX FT New antibody to human collapsin - used to inhibit the activity of
 PT collapsin, to induce neurite out-growth and to treat individuals with
 PT nerve damage.

XX PS Claim 2; Columns 15-18; 11pp; English.

XX CC Aff antibody capable of specifically binding at least a portion of
 CC the collapsin protein can be used to purify human collapsin and
 CC to inhibit the activity of the protein. It can be used to induce
 CC neurite outgrowth by neuronal cells and to treat individuals

CC suffering from nerve damage.

XX SQ Sequence 477 AA;

Alignment Scores:

Pred. No.: 1.18e-227 Length: 477
 Score: 2562.00 Matches: 472
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 3
 Query Match: 92.33% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR74175 (1-477)

QY 899 GAGAGTCACATCTCTGAAGATGACAAAGTATACCTTTCTCCGTGAAAATGCAATAGAT 958
 DB 2 GluHisAspAsnProGluAspAspLysValTrpPhePheArgGluAsnAlaIleAsp 21
 QY 959 GGAGAACACATCTGGAAAAGCTACTCAGCTAGAAATAGGTAGATATGCAAGAATGACTTT 1018
 DB 22 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 41
 QY 1019 GGAGGGCACAGAACTCTGGTGAATAAATGGACAAACATTCCTCAAGCTCTGCTGATTCG 1078
 DB 42 GlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCys 61
 QY 1079 TCAGTGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAACTGCAAGGATGATTCCTA 1138
 DB 62 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 81
 QY 1139 ATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATATGGAGTGTTCAGACTTCCAGTAAC 1198
 DB 82 MetAsnPhelysAspProLysAsnProValValTrpGlyValPheThrSerSerAsn 101
 QY 1199 ATTTTCAAGGATCAGCGTGTGTATGTATAGCATGATGATGAGAGGGTGTTCCTT 1258
 DB 102 IlePheLysGlySerAlaValCysMetTrpSerMetSerAspValArgValPheLeu 121
 QY 1259 GGTCCATATGCCACAGGGATGGACCCCACTATCAATGGGTGCCTTATCAAGGAGAGTGC 1318
 DB 122 GlyProTrpAlaHisArgAspGlyProAsnTrpGlnTrpValProTrpGlnGlyArgVal 141
 QY 1319 CCTATCCACGGCCAGGAACCTTGTCCAGCAAAAACATTTGGTGGTGTTCAGTCTACAAAG 1378
 DB 142 ProTrpProArgProGlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLys 161
 QY 1379 GACTTCTGATGATGATTAACCTTTGCAAGAAGTATCCAGCATGATCAATCCAGTG 1438
 DB 162 AspLeuProAspAspValIleThrPheAlaArgSerHisProAlaMetTrpAsnProVal 181
 QY 1439 TTTCCTATGACAAATCGCCCAATAGTATCAAAACGGATGTAATATTCAATTTTACACAA 1498
 DB 182 PheProMetAsnAsnArgProIleValIleLysThrAspValAsnTrpGlnPheThrGln 201
 QY 1499 ATTGTCGTAGACCGAGTGGATGCAAGAAGATGGACAGATGATGATGTTATGTTTTCGAAACA 1558
 DB 202 IleValValAspArgValAspAlaGluAspGlyGlnTrpAspValMetPheIleGlyThr 221
 QY 1559 GATGTGGACCGTCTTTAAAGTAGTTTCAATTCCTAAGGAGACTGGTGTGATGATTTAGAA 1618
 DB 222 AspValGlyThrValLeuLysValValSerIleProLysGluThrTrpTrpAspLeuGlu 241
 QY 1619 GAGCTTCTGCTGGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGAG 1678
 DB 242 GluValLeuLeuGluGluMetThrValPheArgGluProThrAlaIleSerAlaMetGlu 261
 QY 1679 CTTTCCACTAAGCAGCAACACTATATATTGTTTCAACGGTGGGGTTCGCCAGTCCCT 1738
 DB 262 LeuSerThrLysGlnGlnLeuTrpIleGlySerThrAlaGlyValAlaGlnLeuPro 281
 QY 1739 TTACACCGGTGTGATATTTACGGGAACCGGTGCTGATGTTGCTCGCCCGAGACCTT 1798
 DB 282 LeuHisArgCysAspIleTrpGlyLysAlaCysAlaGluCysCysLeuAlaArgAspPro 301

QY 1799 TACTGTCTGGATGGTTCGATGTTCTCGCTATTTTCCACTGCAAGAGAGCGACA 1858
 Db |||||
 QY 302 TyrCysAlaTrpAspGlySerAlaCysSerArgTyrPheProThrAlaLysArgArgThr 321
 Db |||||
 QY 1859 AGACACCAAGATATAAGAAATGGACACCCACTGACTCTCTGCTTACAGCTTACACCATAT 1918
 Db |||||
 QY 322 ArgArgGlnAspIleArgAsnGlyAspProLeuThrHisCysSerAspLeuHisHisAsp 341
 Db |||||
 QY 1919 AATCACCATGGCCACAGCCCTCAAGAGAGATCATCTATGGTGTAGAGAAATAGTACACA 1978
 Db |||||
 QY 342 AsnHisHisGlyHisSerProGluGluIleTyrGlyValGluAsnSerSerThr 361
 Db |||||
 QY 1979 TTTTGGAAATGACAGTCCGAGTCCGAGAGAGCGCTGTCTATTGGCAATTCCAGAGCGCA 2038
 Db |||||
 QY 362 PheLeuGluCysSerProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArg 381
 Db |||||
 QY 2039 AATGAAGAGCGAAGAGAGATCAGATGAGTGGATGATCATATCATCAGCAGACATCAAGC 2098
 Db |||||
 QY 382 AsnGluGluArgLysGluGluIleArgValAspHisIleIleArgThrAspGlnGly 401
 Db |||||
 QY 2099 CTTCTGCTACGTAGTCTCAACAGAGAGATTCAGGCAATTACCTCTGCGATGCGGTGAA 2158
 Db |||||
 QY 402 LeuLeuLeuArgSerLeuGlnGlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlu 421
 Db |||||
 QY 2159 CATGGTTCATACAACTCTTCTTAAGGTAAACCTGGAAGTCAATTCACAGAGCATTTG 2218
 Db |||||
 QY 422 HisGlyPheIleGlnThrLeuLeuLysValThrLeuGluValIleAspAsnGluHisLeu 441
 Db |||||
 QY 2219 GAAGAACTCTTCATAAGATGATCATGAGATGCTCTTAAGACCAAGAAATGTCGAAT 2278
 Db |||||
 QY 442 GluGluLeuLeuHisLysAspAspGlyHisGlySerLysThrLysGluMetSerAsn 461
 Db |||||
 QY 2279 AGCATGACACCTAGCCAGAAGTCTGTTACAGAGACTTTCATGCGAG 2323
 Db |||||
 QY 462 SerMetThrProSerGlnLysValTrpTyrArgAspPheMetGln 476
 Db |||||

RESULT 6

AA27127
 ID AAY27127 standard; Protein; 777 AA.
 XX
 AC AAY27127;

DT 14-SEP-1999 (first entry)
 XX
 DE Human brain tissue-derived polypeptide (clone OM007).
 XX
 KW Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
 KW recombinant; diagnosis; treatment.
 XX

OS Homo sapiens.
 XX

Key Location/Qualifiers
 FH 1..36
 FT Peptide /note= "signal peptide"
 FT Protein 37..777
 FT /note= "mature protein"

XX WO9933873-A1.
 XX
 XX
 PD 08-JUL-1999.
 XX
 XX 25-DEC-1998; 98WO-JP05952.
 XX
 XX 26-DEC-1997; 97JP-0358811.
 XX
 XX (ON0Y) ONO PHARM CO LTD.
 XX
 XX Fukushima D, Shibayama S, Tada H;
 XX WPI, 1999-419088/35.
 XX
 XX N-PSDB; AAX89112, AAX89113.
 XX

PT New adult human brain tissue-produced polypeptides useful for
 PT diagnosis and treatment
 XX

PS Claim 1; Page 36-39; 86pp; Japanese.
 XX

CC The invention provides polypeptides (AAY27127-Y27133) produced by human
 CC adult brain tissue, human bone marrow or a human umbilical cord venous
 CC endothelial cell. Host cells transformed with vectors comprising the
 CC nucleic acids encoding the polypeptides are used for the recombinant
 CC expression of the polypeptides. The polypeptides can be used in
 CC diagnosis, treatment and basic studies, with wide applications in
 CC treatment depending on the activity to be aimed at. Sequences
 CC AAX89112-125 represent nucleic acids encoding the polypeptides.
 XX

SQ Sequence 777 AA;

Alignment Scores:

Pred. No.: 2,68e-198 Length: 777
 Score: 2245.50 Matches: 413
 Percent Similarity: 73.67% Conservative: 141
 Best Local Similarity: 54.92% Mismatches: 177
 Query Match: 145.86% Indels: 21
 Gaps: 9

US-09-774-490-1 (1-2709) x AAY27127 (1-777)

QY 275 AAGAACAAATGCGCAAGGCTGAATATCTCTACAAAGAAATGTTGGAATCCAAACAATGTG 334
 Db |||||
 QY 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
 Db |||||
 QY 335 ATCACTTTCATGCTGGCCACAGCTCCAGTTATCATATCATCTCTCTTTTGGATGAGAA 394
 Db |||||
 QY 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuAspGluGlu 78
 Db |||||
 QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATTCGACCTGGTTAATATC 454
 Db |||||
 QY 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 Db |||||
 QY 455 ---AAGATTTCACAAAGATGCTGCGCAGTATCTTACACAGAGAGATGAAATGCAAG 511
 Db |||||
 QY 99 AsnLysAsnPheLysLysIleTyrTrpProAlaLysGluArgValGluLeuCysLys 118
 Db |||||
 QY 512 TGGCTGGAAGACATCTCGAAAGATGTCTAATTTTCATCAAGGTACTTAAAGCATAT 571
 Db |||||
 QY 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
 Db |||||
 QY 572 AATCAGACTCATCTGTAGCTGTGGAAACGGGGCTTTTCATCCAAATTTGACCTACAT 631
 Db |||||
 QY 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
 Db |||||
 QY 632 GAAATTGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAC 691
 Db |||||
 QY 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrArgAsnLeuGluSer 178
 Db |||||
 QY 692 GGCGTGGAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAAATAGATGA 751
 Db |||||
 QY 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 Db |||||
 QY 752 GAATTATCTCTGGAAGTCTGAGCTGATTTTATGGGCGAGACTTTGCTATCTCCGAACT 811
 Db |||||
 QY 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
 Db |||||
 QY 812 CTTGGG-----CACCACCCCAATCAGCAGCAGCATGATTTCCAGGTGGCTC 862
 Db |||||
 QY 219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
 Db |||||
 QY 863 AATGATCCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACATCTCTGAAGATGAC 922
 Db |||||
 QY 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAspAsp 258
 Db |||||
 QY 923 AAGTATATCTTTTCTTCCTGCGTAAATGCAATAGATGAGAACACTCTCTGGAAGAGTACT 982
 Db |||||
 QY 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
 Db |||||

QY 983 CAGCTAGATAGTTCAGATATGCAAGAAATGACTTTGGAGGSCACAGAGTCTGGTGAAT 1042
Db
QY 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuLeuIleAsn 298
Db
QY 1043 AAATGACAACTTCCTCAAGCTCGCTGATTGCTCAGTGCAGGTCCAAATGGCATT 1102
Db
QY 299 LysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
Db
QY 1103 GACACTCATTTGATGAAGTGAAGTATGATTCCTTAATCAACTTTAAAGATCCTAAAT 1162
Db
QY 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
Db
QY 1163 CAGTTGTATATGAGTGTTCAGCTCCAGTAAACATTTTCAAGGATCAGCCGTGTGT 1222
Db
QY 339 ProValValTyrGlyValPheThrThrThrSerIlePheLysGlySerAlaValCys 358
Db
QY 1223 ATGTATAGCATGATGTGAAGGGTTCCTTGGTCCATATGCCACAGGATGA 1282
Db
QY 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
Db
QY 1283 CCCAACTCAATGGTGCTTATCAAGGAAGTCCCTATCCAGCGCCAGAACTGT 1342
Db
QY 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
Db
QY 1343 CCCAGCAAAACATTTGGTGT---TTTGACTCTCAAAAGGACTTCCTGATCATGTATA 1399
Db
QY 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
Db
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGAACATCCCA 1459
Db
QY 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
Db
QY 1460 ATAGTATCAAAACGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db
QY 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
Db
QY 1520 CGAGAGATGGACAGTATCATGTTATGTTTATCGGAACAGATTTGGGACCGTCTTAA 1579
Db
QY 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
Db
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db
QY 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
Db
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACATAGCAGCAACA 1699
Db
QY 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
Db
QY 1700 CTATATATTGGTTCAACGGCTGGGTTCGCCAGCTCCCTTACACGGTGTGATATTAC 1759
Db
QY 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
Db
QY 1760 GGAAGACGCTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTTGGATGTTCT 1819
Db
QY 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
Db
QY 1820 GCATGTTCTCGCTATTTTCCCTGCAAGAGAGCCGACAGCAGCAGATATAGAAT 1879
Db
QY 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValIysTyr 577
Db
QY 1880 GGAGACCCACTGACTCACCTGTTCCAGCTTACACCATGATTAATCACCATGCCACGCCCT 1939
Db
QY 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GlnThrAla 596
Db
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTAGCACATTTTGAATGCAAGTCCGAAG 1999
Db
QY 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
Db
QY 2000 TGCAGAGAGCGTGTGCTATTGGCAATTCAGAGCGCAATTAAGAGCGCAAAAGAG 2059
Db
QY 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGlu 636
Db

QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTCTAGTAGTGTACAA 2119
Db
QY 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
Db
QY 2120 CAGAAGGATTCAGCAATTACCTTCGCATCCGCTGGAAACATGGGTTTCATACAACTCTT 2179
Db
QY 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
Db
QY 2180 CTTAAGGTAAACCTGGAGTCAATTCACACAGAGCATTTGGGAAGAACTTCTTCATAAAGAT 2239
Db
QY 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
Db
QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db
QY 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
Db
QY 2300 GTCTGTACAGAGACTTTCATCAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db
QY 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
Db
QY 2360 TTCTGTGACAACTTTGGAAAGGACCGCAAAACACGTCGCAAGGCCAGGACATACC 2419
Db
QY 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
Db
QY 2420 CCAGGAACACTAAACAAATGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db
QY 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
Db
QY 2480 ACCCAC---GAATTTGAGAGGCGCCACAGGAGTGTCT 2512
Db
QY 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
Db
RESULT 7
AAAY99427
ID AAAY99427 standard; Protein; 777 AA.
XX AC AAAY99427;
XX DT 08-AUG-2000 (first entry)
XX DE Human PRO1491 (UNQ760) amino acid sequence SEQ ID NO:310.
XX KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
XX KW transmembrane; secretion; immunoadhesion; pharmacological; screening.
XX OS Homo sapiens.
XX PN WO200012708-A2.
XX PD 09-MAR-2000.
XX PF 01-SEP-1999; 99WO-US20111.
XX PR 01-SEP-1998; 98US-0098716.
XX PR 01-SEP-1998; 98US-0098749.
XX PR 01-SEP-1998; 98US-0098750.
XX PR 02-SEP-1998; 98US-0098803.
XX PR 02-SEP-1998; 98US-0098821.
XX PR 02-SEP-1998; 98US-0098843.
XX PR 09-SEP-1998; 98US-0099536.
XX PR 09-SEP-1998; 98US-0099596.
XX PR 09-SEP-1998; 98US-0099598.
XX PR 09-SEP-1998; 98US-0099602.
XX PR 09-SEP-1998; 98US-0099642.
XX PR 10-SEP-1998; 98US-0099741.
XX PR 10-SEP-1998; 98US-0099754.
XX PR 10-SEP-1998; 98US-0099763.
XX PR 10-SEP-1998; 98US-0099792.
XX PR 10-SEP-1998; 98US-0099808.
XX PR 10-SEP-1998; 98US-0099812.
XX PR 10-SEP-1998; 98US-0099815.
XX PR 10-SEP-1998; 98US-0099816.
XX PR 15-SEP-1998; 98US-0100385.

```
PR 15-SEP-1998; 98US-0100388;
PR 15-SEP-1998; 98US-0100390;
PR 16-SEP-1998; 98US-0100584;
PR 16-SEP-1998; 98US-0100627;
PR 16-SEP-1998; 98US-0100661;
PR 16-SEP-1998; 98US-0100662;
PR 16-SEP-1998; 98US-0100664;
PR 17-SEP-1998; 98US-0100683;
PR 17-SEP-1998; 98US-0100684;
PR 17-SEP-1998; 98US-0100711;
PR 17-SEP-1998; 98US-0100711;
PR 17-SEP-1998; 98US-0100930;
PR 18-SEP-1998; 98US-0100848;
PR 18-SEP-1998; 98US-0100849;
PR 18-SEP-1998; 98US-0101014;
PR 18-SEP-1998; 98US-0101068;
PR 18-SEP-1998; 98US-0101071;
PR 22-SEP-1998; 98US-0101279;
PR 23-SEP-1998; 98US-0101471;
PR 23-SEP-1998; 98US-0101472;
PR 23-SEP-1998; 98US-0101474;
PR 23-SEP-1998; 98US-0101475;
PR 23-SEP-1998; 98US-0101476;
PR 23-SEP-1998; 98US-0101477;
PR 23-SEP-1998; 98US-0101479;
PR 24-SEP-1998; 98US-0101738;
PR 24-SEP-1998; 98US-0101741;
PR 24-SEP-1998; 98US-0101743;
PR 24-SEP-1998; 98US-0101915;
PR 24-SEP-1998; 98US-0101916;
PR 29-SEP-1998; 98US-0102207;
PR 29-SEP-1998; 98US-0102240;
PR 29-SEP-1998; 98US-0102307;
PR 29-SEP-1998; 98US-0102330;
PR 29-SEP-1998; 98US-0102331;
PR 30-SEP-1998; 98US-0102484;
PR 30-SEP-1998; 98US-0102487;
PR 30-SEP-1998; 98US-0102570;
PR 30-SEP-1998; 98US-0102571;
PR 01-OCT-1998; 98US-0102684;
PR 01-OCT-1998; 98US-0102687;
PR 02-OCT-1998; 98US-0102965;
PR 06-OCT-1998; 98US-0103258;
PR 06-OCT-1998; 98US-0103449;
PR 07-OCT-1998; 98US-0103314;
PR 07-OCT-1998; 98US-0103315;
PR 07-OCT-1998; 98US-0103328;
PR 07-OCT-1998; 98US-0103395;
PR 07-OCT-1998; 98US-0103396;
PR 07-OCT-1998; 98US-0103401;
PR 08-OCT-1998; 98US-0103633;
PR 08-OCT-1998; 98US-0103678;
PR 08-OCT-1998; 98US-0103679;
PR 08-OCT-1998; 98US-0103711;
PR 14-OCT-1998; 98US-0104257;
PR 20-OCT-1998; 98US-0104987;
PR 20-OCT-1998; 98US-0105000;
PR 21-OCT-1998; 98US-0105002;
PR 21-OCT-1998; 98US-0105104;
PR 22-OCT-1998; 98US-0105169;
PR 22-OCT-1998; 98US-0105266;
PR 26-OCT-1998; 98US-0105693;
PR 26-OCT-1998; 98US-0105694;
PR 27-OCT-1998; 98US-0105807;
PR 27-OCT-1998; 98US-0105881;
PR 27-OCT-1998; 98US-0105882;
PR 28-OCT-1998; 98US-0106062;
PR 28-OCT-1998; 98US-0106023;
PR 28-OCT-1998; 98US-0106029;
PR 28-OCT-1998; 98US-0106030;
PR 28-OCT-1998; 98US-0106032;
PR 28-OCT-1998; 98US-0106033;

PR 28-OCT-1998; 98US-0106178;
PR 29-OCT-1998; 98US-0106248;
PR 29-OCT-1998; 98US-0106384;
PR 29-OCT-1998; 98US-0108500;
PR 30-OCT-1998; 98US-0106464;
PR 03-NOV-1998; 98US-0106856;
PR 03-NOV-1998; 98US-0106902;
PR 03-NOV-1998; 98US-0106905;
PR 03-NOV-1998; 98US-0106919;
PR 03-NOV-1998; 98US-0106932;
PR 03-NOV-1998; 98US-0106934;
PR 10-NOV-1998; 98US-0107783;
PR 17-NOV-1998; 98US-0108775;
PR 17-NOV-1998; 98US-0108779;
PR 17-NOV-1998; 98US-0108787;
PR 17-NOV-1998; 98US-0108788;
PR 17-NOV-1998; 98US-0108801;
PR 17-NOV-1998; 98US-0108802;
PR 17-NOV-1998; 98US-0108806;
PR 17-NOV-1998; 98US-0108807;
PR 17-NOV-1998; 98US-0108867;
PR 17-NOV-1998; 98US-0108925;
PR 18-NOV-1998; 98US-0108848;
PR 18-NOV-1998; 98US-0108849;
PR 18-NOV-1998; 98US-0108850;
PR 18-NOV-1998; 98US-0108851;
PR 18-NOV-1998; 98US-0108852;
PR 18-NOV-1998; 98US-0108858;
PR 18-NOV-1998; 98US-0108904;
XX
XX (GETH ) GENENTECH INC.
XX
PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
XX N-PSDB; AAA37109.
XX
PT New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
XX small molecule inhibitors of the relevant receptor/ligand interactions
XX
PS Claim 12; Fig 176; 773pp; English.
XX
CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AA999340 to AA999462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding then have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
SQ Sequence 777 AA;

Alignment Scores:
Pred. No.: 2,68e-198 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 21 Gaps: 9

US-09-774-490-1 (1-2709) x AA999427 (1-777)

Qy 275 AAGAACAAATGTGCAAGCTGAATATCTACAAAGAAATGTTGGATCAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
Qy 335 ATCATTTCATGGCTTGGCCACACAGCTCCAGTTATCATCTCTCTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
```

QY	395	CGGAGTAGCGTGATGATGTGGAGCAAAAGATACACATATTTTCATTCGACCTGGTTATATATC	454
DB	79	ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu	98
QY	455	---AAGATTTTCAAAGATTGTCGGCCAGTATCTTACACGACGACAGATGAATCGAAG	511
DB	99	AsnLysAsnPheTyrLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys	118
QY	512	TGGCTCGAAAAGACATCCTGAAAGAATGTCTAAATTTTCATCAAGGTACTTTAAGGCATAT	571
DB	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr	138
QY	572	AATCAGACTCATCTTGATCGCCTGTGGAAACGGGGCTTTTCATCAAATTTGCACCTACATT	631
DB	139	AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle	158
QY	632	GAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAC	691
DB	159	AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer	178
QY	692	GGCCGTGGGAGAGTCCATATGACCCCTAAAGCTGTGACAGCATCCCTTTTAAATAGATGGA	751
DB	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu	198
QY	752	GAATTATATCTCGAACTGCAGCTGATTTTATGGGGCGAGACTTTGTCTATCTCCGAACT	811
DB	199	TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer	218
QY	812	CTTGGG-----CACCACCACCAATCAGGACAGACGACGATGATTCACAGTGGCTC	862
DB	219	LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu	238
QY	863	AATGATCCAAAGTTTCATTAGTGGCCACTCATCTCAGAGAGTGACATCTCGAAGAAGCTACT	922
DB	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp	258
QY	923	AAAGTATATCTTTTCTCCGTGAAAATCGAATAGATGGAGAACTCTGAAAAGCTACT	982
DB	259	LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle	278
QY	983	CACGCTAGATAGTTCAGATATGCAAGATGACTTTTCGAGGGCACGAAGTCTGTGTGAAT	1042
DB	279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn	298
QY	1043	AAATGGACAACATTCCTCAAAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAATGGCATT	1102
DB	299	LysTrpThrThrPheLeuLysAlaArgLeuLecysSerIleProGlySerAspGlyAla	318
QY	1103	GACACTCATTTTGATGAATCGCAGAGTGATTCCTCAATGAACCTTTAAAGATCTCAAAAAT	1162
DB	319	AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn	338
QY	1163	CCAGTTGTATATGAGTGTTTACACTTCGAGTAAACATTTTCAAGGAGTACAGCGGTGT	1222
DB	339	ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys	358
QY	1223	ATGTATAGCATGATGTGCAGAGGTGTTCCTTCGTCCATATGCCACAGGATGGA	1282
DB	359	ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer	378
QY	1283	CCCAACTATCAATGGGTGCCTTATCAAGAGAGAGTCCCTCATCCACGGCCAGAACTTGT	1342
DB	379	AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys	398
QY	1343	CCGAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCCTGATGATGTATA	1399
DB	399	ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle	418
QY	1400	ACCTTTGCAAGAAGTTCATCCAGCCATGTACAATCCAGTGTTCCTCATGAACAATCGCCA	1459
DB	419	SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro	438
QY	1469	ATAGTGATCAAAACGGAGTGAATTTATCAATTTACACAAATTTGCTAGACCGAGTGGAT	1519

Db	439	ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle	458
Qy	1520	GCACGAAGATGCACAGTAGTATGATGTTATCGGAACACAGATCTGGGACCGCTTCTTAA	1579
Db	459	AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys	478
Qy	1580	GTACTTTCAATTCCTAAGAGACACTGTGATATTAGAGAGGTTCTGCTGGAGAAATG	1639
Db	479	ValValSerIleSerLysGluLysTrp--AsnMetGluGluValValLeuGluGluLeu	497
Qy	1640	ACAGTTTTCGGNAACCGACTGCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA	1699
Db	498	GlnIlePheLysHisSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln	517
Qy	1700	CTATATATTGTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTGATATTTAC	1759
Db	518	LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr	537
Qy	1760	GGGAAAGCGTGTGCTAGTGTGCTCGCCGAGACCTTACTGTGCTCTGGGATGTTCT	1819
Db	538	GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn	557
Qy	1820	GCATGTTCTCGCTATTTTCCACTGCAAGAGAGCACAAGACAGCATATAAGAAT	1879
Db	558	AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr	577
Qy	1880	GGACACCACTGACTCACTGTCAGACTTACACCATGATAATCACCATGCCACACGCCCT	1939
Db	578	GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis--GluThrAla	596
Qy	1940	GAAGAGAGAATCATCTATGTGTAGAGAAATAGTAGCACATTTTGGAAATGCAGTCCGAAG	1999
Db	597	AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys	616
Qy	2000	TCGCAGAGACGCTGTCTATTGTCGAATTCACAGGCGGAAATCAAGACGCGAAAGAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu	636
Qy	2060	ATCAGATGGATGATCATATCATCAGACAGATCAAGGCCCTCTGCTAGCTAGTCTACAA	2119
Db	637	LeuLysProAspGluArgIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAGAAGATTCCAGCAATTACCTCTGCATCGCGTGGAAACATGGTGTTCATAAACTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTGGGAAGTTCATTGACACAGACGATTTGGAAGAATCTTCTCATAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAA	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGAGACTTCATGCAGCTCATCAACACCCCAATCTCAACAGCATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe--SerLeuAspGln	729
Qy	2360	TTCTGTGAACAAGTTTGGAAAGGACCGGAAACACAGCTCGCAAGGCCGAGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGGAACGATACAAATGGAACGACTTACAGAAATTAAGAAAGCTAGAAACAGGAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACCCAC---GAATTGTAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775

RESULT 8

AAU29197

Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrThrAsnProAspAsp 258
QY 923 AAAGTATACCTTTCTTCGCGTAAATGCAATAGATGGAGACACTCTGGAAGCTACT 982
Db 259 LysIleTyrPhePhePheArgGluSerGlnGlySerThrSerAspLysThrIle 278
QY 983 CAGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGGCCACAGAGTCTGCTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyClnArgSerLeuIleAsn 298
QY 1043 AAATGGACAACATTCCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCATT 1102
Db 299 LysTyrThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTGATGAACCTGAGGATGATTCTTCAATCACTTAAGATCCTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATATGGAGTGTTCAGACTTCAGTAAACATTTTCAAGGGATCAGCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGATGATGAGAGGGTGTCTCTGTTGCTCCATATGCCACAGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTATCAATGGTGTCTTATCAGGAAGTCCCTATCCAGGCCAGGAATGTTGT 1342
Db 379 AlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTTGGTGGT--TTTGACTCTACAAAGGAGCTTCTGATCATGTATATA 1399
Db 399 ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGTAATCAATTCACAAATTCGTAGACCGAGTGAT 1519
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACGGATTAATTAATCAATTCACAAATTCGTAGACCGAGTGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 CGAGAAGATGGACATGATGTTATGTTTATCGGAACAGATGTTGGACCGCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGTATGATATTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTyr--AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTTCGGGACCGACTGCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACCGGTTGATATTATAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGGGATGCTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaIleAspGlyAsn 557
QY 1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACCCACAGACGACAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyr 577
QY 1880 GGAGACCCCACTCACTCACTGTTTCAGACTTACCATGATAATCACCATGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTyrAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGCTAGAGATATAGTACATATTTTGAATGCTAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616

QY 2000 TCGCAGAGACGCTGGTCTATTGGCAATTCCAGAGCGCAATTAAGAGCGCAAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTyrTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGSCCTTCTGCTACGTAGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGATTTCAGCAATTACCTCTGCCATCGGTGGAAACATGGTTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAGTTCATGACACAGCATTTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGGCTCTTAAGACCAAAATAATGTCATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGACGCTCATCAACACCCCAATCTCAACACGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAGGACCGCAAAACAAACGTCGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTyrPheArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACATAACAATGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTyrLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCAC---GAATTTGAGAGGCGCACCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 9

AA866176
ID AA866176 standard; protein; 777 AA.
AC AA866176;
XX
DT 02-APR-2001 (first entry)
XX
DE Protein of the invention #88.
XX
KW Secreted; transmembrane; gene therapy.
XX
OS Unidentified.
XX
PN WO200078961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US04342.
XX
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 30-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
XX
XX (GETH) GENENTECH INC.

Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard AF, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;

PI Watanabe CK, Williams PM, Wood WI;
 XX WPI; 2001-071395/08.
 XX
 XX Secreted and transmembrane proteins and nucleic acids designated PRO,
 PT useful as hybridization probes, in chromosome and gene mapping and gene
 PT therapy -
 XX
 XX Claim 1; Fig 176; 787pp; English.
 PS
 XX The present invention relates to secreted and transmembrane proteins.
 CC These proteins and the DNA encoding them may be used as hybridization
 CC probes, in chromosome and gene mapping and in the generation of
 CC anti-sense RNA and DNA. They may also be used to generate either
 CC transgenic animals or knockout animals which are in turn useful for
 CC development and screening of therapeutically useful reagents.
 CC The nucleic acids may also be used in gene therapy.
 XX
 SQ Sequence 777 AA;

Alignment Scores:
 Pred. No.: 2,686-198 Length: 777
 Score: 2245.50 Matches: 413
 Percent Similarity: 73.67% Conservative: 141
 Best Local Similarity: 54.92% Mismatches: 177
 Query Match: 45.86% Indels: 21
 DB: 22 Gaps: 9

US-09-774-490-1 (1-2709) x AAB66176 (1-777)

QY 275 AGAACATGTCGCAAGGCTGAATATTCCTCAAGAAATGTTGGATCCAACTGTG 334
 DB LysGlnAsnIleProArgLeuLysLeuThrTyLysAspLeuLeuLeuSerAsnSerCys 58
 QY 335 ATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTCTCTTTTGGATGAGGA 394
 DB IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu 78
 QY 395 CGAGTAGGCTGTATGTTGGACAAAGGATCATATTTTCATTCGACCTGGTTAATATC 454
 DB ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 QY 455 ---AGGATTTCAAGATGTTGGCCAGTATCTTACACAGAAAGATGAATGAAG 511
 DB AsnLysAsnPheLeuLysIleTyTrpProAlaLysGluArgValGluLeuLys 118
 QY 512 TGGGCTGGAAAGACATCCTGAAGAATGCTGAATTTTCATCAAGTACTTAAGCCATAT 571
 DB LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTy 138
 QY 572 AATCAGACTCACTGTAGCCCTGTGGAAACGGGGCTTTTCATCCAAATTTGCACCTACAT 631
 DB AsnLysThrHisIleTyValCysGlyThrGlyAlaPheHisProIleCysGlyTyTrile 158
 QY 632 GAAATGACATCTCCGAGACAAATATTTTAAAGTGGAGAACTACATTTTGAANAAC 691
 DB AspLeuGlyValTyLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
 QY 692 GSCCGTGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
 DB GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 QY 752 GAAATATATCTGGAACCTGACGCTGATTTTATGGGGCGAGACTTTTGTCTATCTCCGAAC 811
 DB TyrLeuTySerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
 QY 812 CTGGG-----CACACACCCATCAGGACAGACGATGATTCAGGTTGGCTC 862
 DB LeuGlyProThrHisAspHisTyIleArgThrAspIleSerGluHisTyTrileu 238
 QY 863 AATGATCCAAAGTTCATTAGTCCCACTCTCATCTCAGAGAGTGACATCTCTGAGATGAC 922
 DB AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyAsnProAspAsp 258

QY 923 AAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTCTGAAAGCTACT 982
 DB LysIleTyPhePhePheArgGluSerGlnGluGlySerThrSerAspLysThrile 278
 QY 983 CACGCTAGAATAGTCAGATATGCAAGATGACATTTGGAGGCGACAGAGTCTGGTGAAT 1042
 DB LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
 QY 1043 AAATGGCAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCGAGGTCCTCAATGGCAT 1102
 DB LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
 QY 1103 GACACTCAATTTTGAATGACACTGCAGGATGTATTCCTAATGAATTTTAAAGATCTCT 1162
 DB AspThrTyPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAspGluArgAsn 338
 QY 1163 CCAGTTGATATGAGTGTGTTTACGACTTCAGTCCAGTAACTTTTCAAGGGATGACCGTGT 1222
 DB ProValValTyGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
 QY 1223 ATGTATAGCATGATGTGAGAAAGGTGTTCTTGTGTCATATGCCACAGGATGGA 1282
 DB ValTySerMetAlaAspIleArgAlaValPheAsnGlyProTyAlaHisLysGluSer 378
 QY 1283 CCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCAGCGCCAGGAACTTGT 1342
 DB AlaAspHisArgTrpValGlnTyAspGlyArgIleProTyProArgProGlyThrCys 398
 QY 1343 CCCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTCTGATGATGTATA 1399
 DB ProSerLysThrTyAspProLeuIleLysSerThrArgAspPheProAspAspValile 418
 QY 1400 ACCTTTCAGAAAGTTCATCCAGCCATGTACAATCCAGTGTCTTCTATGAAACAATCGCCA 1459
 DB SerPheLysArgHisSerValMetTyLysSerValTyProValAlaGlyGlyPro 438
 QY 1460 ATAGTGTCAAAACCGGATGTAATTTATCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
 DB ThrPheLysArgIleAsnValAspTyArgLeuThrGlnIleValValAspHisValile 458
 QY 1520 GCAGAAATGACAGTATGATGTTATGTTATCCGAACAGATGTTGGACCGTCTCTTAAA 1579
 DB AlaGluAspGlyGlnTyAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
 QY 1580 GTAGTTTCAATCTCAAGGAGACTTGTGTATGATTTAGAAGAGTCTCTCGTGGAGAATG 1639
 DB ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
 QY 1640 ACAGTTTTCGGGAACCGCATGCTATTTTCAGCAATGAGCTTTCCACTTAAGCAGCAACA 1699
 DB GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
 QY 1700 CTATATATGTTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
 DB LeuTyIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTy 537
 QY 1760 GGGAAACGGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGCTGCTGGATGGTGTCT 1819
 DB GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyCysAlaTrpAspGlyAsn 557
 QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGACAGATATATGAANAAT 1879
 DB AlaCysSerArgTyAlaProThrSerLysArgAlaArgAGGlnAspValLysTy 577
 QY 1880 GGAGACCCACTGACTCTGTTGAGTTCACCATCATATCAATCAATGCGCCACAGCCCT 1939
 DB GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
 QY 1940 GAAGAGAGATCATCTATGCTGTAGACAATAGTAGACACATTTTGGATCCAGTCCCAAG 1999
 DB AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616

PR	18-AUG-1998;	98US-096949P.
PR	18-AUG-1998;	98US-096959P.
PR	18-AUG-1998;	98US-097022P.
PR	26-AUG-1998;	98US-097952P.
PR	26-AUG-1998;	98US-097954P.
PR	26-AUG-1998;	98US-097955P.
PR	26-AUG-1998;	98US-097971P.
PR	26-AUG-1998;	98US-097974P.
PR	01-SEP-1998;	98US-098716P.
PR	02-SEP-1998;	98US-098723P.
PR	02-SEP-1998;	98US-098803P.
PR	02-SEP-1998;	98US-098821P.
PR	09-SEP-1998;	98US-098843P.
PR	10-SEP-1998;	98US-099741P.
PR	10-SEP-1998;	98US-099754P.
PR	10-SEP-1998;	98US-099763P.
PR	10-SEP-1998;	98US-099812P.
Alignment Scores:		
Pred. No.:	2,68e-198	Length: 777
Score:	2245.50	Matches: 413
Percent Similarity:	73.67%	Conservative: 141
Best Local Similarity:	54.92%	Mismatches: 177
Query Match:	45.86%	Indels: 21
DB:	24	Gaps: 9
US-09-774-490-1 (1-2709) x ABU71285 (1-777)		
Qy	275	AAGAACAATGTGCAAGCTGAATAATTCCTTACAAAGAAATGTTGGAATCCCAACAATGTG 334
Dd	39	LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
Qy	335	ATCACATTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTCCTCTTTGGATGAGAA 394
Dd	59	IleProPheLeuGlySerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
Qy	395	CGGAGTAGCTGTATGTTGGAGCAAAGGATCACATATTTTCATTCGACCTGGTTAATATC 454
Dd	79	ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
Qy	455	---AAGATTTCAAAAGATTGTGGCCAGTATCTTACCAGAGAGAGATGAATGCAGAG 511
Dd	99	AsnLysAsnPhelLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
Qy	512	TGGCTCGAANAAGACATCCTGAAAGATGCTGTAATTTTCATCATCAGGTACTTAAGCATAT 571
Dd	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
Qy	572	AATCAGACTCACATTGTACGCTGTGGAACGGGGCTTTTCATCCAATTTGCCACCTACATT 631
Dd	139	AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
Qy	632	GAAATTGGACATCATCCTCGAGGACAAATATTTTTAAAGCTGGGAACATCACATTTTGAAC 691
Dd	159	AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
Qy	692	GGCCGTGGGAGAGTCCATATGACCCCTAAGCTGCTGACAGATCCCTTTTATAGATGGA 751
Dd	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
Qy	752	GAATTATCTCTGMACTGCAGCTGATTTTATGGGCGAGAGCTTTGCTATCTTCGCACT 811
Dd	199	TyrLeutyrseryGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
Qy	812	CTTTGGG-----CACCACCACCAATCATCAGGACAGAGCATGATTCCAGGTGGCTC 862
Dd	219	LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
Qy	863	ATGATCCAAAGTTCATTAGTCGCCACCTCATCTCAGAGAGTGACAATCCTGAAAGATGAC 922
Dd	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258

QY	923	AAAGTATATCTTTTCTTCCTCGTGAATAATGAGTGGAGAACACTCTCGAAAGAGCTTACT	982
DB	259	LYSILETYRPHEPHEAARGLUSERSERGLNGLUGLYSETHRSERASPLYSTHRILLE	278
QY	983	CACGCTAGAAATAGTGTACAGATATGCAAGAATGACTTTTGAGGAGCACAGAAGTGTGGTGAAT	1042
DB	279	LEUSERGVALGIVATGVALCYLYSASNASPVALGLYGLYGLINARGSERLEULEASN	298
QY	1043	AAATGGACAACATTCCTCAAGCTCGTGTGATTTGCTCAGTGGCCAGGTCCAAATGGCAAT	1102
DB	299	LYSTRPThrPHELEULYSALARGLEULECYSSERILEProGLYSERASPGLYALA	318
QY	1103	GACACTCATTTTGTAGAACTCCAGATGTATTCCTAATGAACCTTAAAGACTCCTAAAAAT	1162
DB	319	ASPTHrTYRPHASPGLULEULNASPILEYrLEULEUPrOTHrARGASPGLUARGASN	338
QY	1163	CCAGTTGTATAGGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTGT	1222
DB	339	PROVALValTYRGlyVALPHEThrThrSERSerILEPHELYSGLYSERALAVAlCYs	358
QY	1223	ATGTATAGCATGAGTGTCTCAGAAAGGTGTTCCTTGGTCCATATGCCACAGGATCGGA	1282
DB	359	VALTYRSerMETAlAASPILeARGALAVAlPHEASNGLYProTYrAlAHISLYSGLUSeR	378
QY	1283	CCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGAACCTTGT	1342
DB	379	ALASPHISARGTTPVALGINTYrASPGLYARGILEProTYrPROARGProGLYThrCYs	398
QY	1343	CCCAGCAAAACATTTGGTGGT---TTTGACTCTACAAGGACCTTCCTGTATGATGTATA	1399
DB	399	PROSErLYSThrTYRASPProLEULeLYSSerThrARGASPHepHEProASPASPVALILE	418
QY	1400	ACCTTTGCCAAGAAGTCATCCAGCCATGTACAACTCCAGTGTTCCTATGAACAATCGCCCA	1459
DB	419	SERPhEILELYsARGHISerVALMetTYrLYSSerVALTYrPROVALAlaGLYGLYPro	438
QY	1460	ATAGTGATCAAAACGGAGTGAATTTATCAATTTACAAAAATTGCGTAGACCAGTGGAT	1519
DB	439	ThrPHELYsARGILEASNVALSPSYrARGLeUTHrGLNILEVALValASPHISValILE	458
QY	1520	GCAGAAGATGCACATGATGATGTTATGTTATCGGAACACAGATGTGGGACCGTTCCTAAA	1579
DB	459	AlaGLUASPGLYGLINTYrASPVAlMetPHELeUGLYThrASPILeGLYThrVALLeULYS	478
QY	1580	GTAGTTTCAATTCTAAGGACACTTGGTATGATTTAGAAGAGTTCTGCTGGGAAGAAATG	1639
DB	479	VALValSERILESerLYsGLULYsTrp---ASNMetGLUGLUValValLeUGLUGLULeU	497
QY	1640	ACAGTTTTTCCGGGAACCGACTGCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA	1699
DB	498	GLNILEPHELYsHISerSERILELeULEUASNMetGLULeUSERLeULYSGLINGLIn	517
QY	1700	CTATATATTGGTTCAACCGCTGGGGTGTGCCAGACTCCCTTTTACACCGGTGTGATATTAC	1759
DB	518	LeUTYrILEGLYSerARGASPGLYLeUVALGLNLeUSERLeUHISARGCYsASPTHrTYr	537
QY	1760	GGGAAAGCGTGTCTGAGTGTGTCTCGCCGAGACCCCTTACTGTGTTGGAGTGGTCT	1819
DB	538	GLYLYsAlACYSAlAASPcYSLeUALeARGASPHProTYrCYsAlATrPASPGLYASN	557
QY	1820	GCATGTCTCCGCTATTTTCCACATGCAAGAGAGCGCACAGACAGATATAAGAAT	1879
DB	558	AlACYSerARGTYrAlAProThrSERLYsARGALaARGARGGLNASPVVALYSer	577
QY	1880	GGAGACCCACTGACTCTGTTAGACTTACCATGATGATATCACCATGCCACAGCCCT	1939
DB	578	GLYASPProILEThrGLINCYSTIPASPILeGLUASPserILESerHIS---GLUTHrAlA	596
QY	1940	GAAGAGAGATCATCTATGTTGTGATAGATAGTAGACACATTTTGGAAATCGACTCGAAG	1999
DB	597	ASPGLULYSVALILEPHEGLYILEGLUPHEASNserThrPHELeUGLYCYsILEProLYs	616

QY	2000	TCGCAGAGAGCGTGCTATTATTCGCAATTCAGAGGCGAAATGAGAGCGCAAAGAAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTyrTrpIleGlnArgSerGlyAspGluHisArgGluGlu	636
QY	2060	ATCAGAGTGATGATCATCATCATCAGGACAGATCAAGGCTTCCTGCTAGCTAGCTACAA	2119
Db	637	LeuLysProAspGluArgIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
QY	2120	CAGAAGGATTTCAGCAATACCTCTCGCATCGCGTGGAAATGGGTTTCATACAACTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
QY	2180	CTTAAGGTAAACCTCGGAAGTCATGACACAGACATTTGGAAGAACTTCTTCATAAGAT	2239
Db	677	VallLysLeuThrLeuAsnValIleGluAenGluGlnMetGluAsnThrGlnArgAlaGlu	696
QY	2240	GATGATGGAGATGGCTCTTAAGACCAAGAATAATGTCANTAGCATGACACCTACCCAGAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
QY	2300	GTCTGGTACAGAGACTTCATGCTAGCTCATCAACCCACCCCAATCTCAACAGATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
QY	2360	TTCTGTGAACAAGTTGGTGGAAAGGACCGCAAAACAACGTGCGCAAAAGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
QY	2420	CCAGGGAAACAGTAAACAATGGAAGCACTTACAGAANAATGAAGAAGGTAGAAACAGGAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
QY	2480	ACCCAC---CAATTTCAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 11			
ID	ABU65742	standard; Protein; 777 AA.	
XX	AC	ABU65742;	
XX	DT	19-MAY-2003 (first entry)	
XX	DE	Human secreted/transmembrane protein, SEQ ID 348.	
KW	KW	Human; PRO; secreted protein; transmembrane protein;	
KW	KW	cytostatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;	
KW	KW	colon tumour; breast tumour; prostate tumour; rectal tumour;	
KW	KW	cervical tumour; liver tumour; TNF-alpha release; arthritis;	
KW	KW	tumour necrosis factor alpha; chondrocyte cell; bone disorder;	
XX	XX	cartilage disorder; sports injury.	
XX	OS	homo sapiens.	
XX	PN	US2003036156-A1.	
XX	XX	20-FEB-2003.	
PD	XX	02-JUL-2002; 2002US-0188767.	
PF	XX	16-SEP-1998; 98WO-US19330.	
PR	PR	07-OCT-1998; 98WO-US21141.	
PR	PR	01-DEC-1998; 98WO-US25108.	
PR	PR	08-MAR-1999; 99WO-US05028.	
PR	PR	14-MAY-1999; 99WO-US10733.	
PR	PR	02-JUN-1999; 99WO-US12252.	
PR	PR	01-SEP-1999; 99WO-US20111.	
PR	PR	15-SEP-1999; 99WO-US21090.	
PR	PR	01-DEC-1999; 99WO-US28301.	
PR	PR	02-DEC-1999; 99WO-US28551.	
PR	PR	30-DEC-1999; 99WO-US31274.	
PR	PR	05-JAN-2000; 2000WO-US00219.	

PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.

PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086021P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088325P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.

```

PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096881P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Alignment Scores:
Pred. No.: 2,686-198 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
Db: 24 Gaps: 9

US-09-774-490-1 (1-2709) x ABU65742 (1-777)

QY 275 AAGAACATGTCAGGCTGAAATATCTCTACAAAGAAATGTTGGAATCCAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGAA 394
Db 59 IleProLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu 78
QY 395 CGAGTAGCTGTATGTGGACCAAGGATCATATTTTCATTCGACTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTGGAAGACATCTGAAAGATGTGCTAATTTTCATCAAGGTACTTAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 ATCAAGCTCACTTGTAGCCCTGTGGACGGGGCTTTTCATCCCAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGGACATCATCTCAGGACAAATATTTTAAAGCTGGAGAACTCACATTTGAAAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCCCGTGGGAAGAGTCCATATGACCTTAAGCTGTGCACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAACCTGATTTTATGGGGCGAGACTTTGCTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCACCAATCAGGACAGCAGCATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCCTCGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258

```

```

QY 923 AAGTATACTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGACTACT 982
Db 259 LysIleTyrPhePheArgGluSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CACGCTAGAATAGCTCAGATATGCAAGAAATGACTTTTGGAGGCGACAGAACTGTGTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
QY 1043 AATGGACAACATTCCTCAAGAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAAATGGCATT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTTGTGAAGTGCAGAGTGTATTCCTAATGAACCTTTAAAGATCCTTAAAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATATGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGATGATGAGAGAGGTGTTCTTGTCTCATATGCCACAGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTTATCCAGGCCCAGGAACTTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGAGCCTTCTCTGATGATGTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAAGTCTACGCCATGTACAAATCCAGTGTTCCTATCAACAATCCGCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAAAATTTTACAAAATTTGCTGAGACCGAGTGCAT 1519
Db 439 ThrPheLysArgIleAsnValAspIlyArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAAGATGACAGTATGATGTTATGTTATCGAAACAGATGTTGGAGCCGTCTTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCGAGAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGAGCTTTCACCTTAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATGTTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTATAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTGTGAGTGTGTGCTCGCCGAGACCTTACTGTCTGGATGGTCTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCACAGACGACAAAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCATCTGTTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGAAATCATCTGTTGTAGAGAAATAGTAGCACATTTTGGATTCAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616

```

QY 2000 TCGCAGAGCGCTGCTATTGGCAATTCAGAGCGCAATGAAGACCGCAAAAGAGAG 2059
 Db 617 SerGlnGlnAlaThrIleLeuTyrTrpTyrIleGlnArgSerGlyUHIleArgGluGln 636
 QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCTTCTGCTAGCTAGCTACAA 2119
 Db 637 LeuLysProAspGluArgIleLeuLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
 QY 2120 CAGAGGATTCAGCAATTAACCTTCGTCATCGGTGGAACTGGGTTTCATCAAACTCTT 2179
 Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
 QY 2180 CTTAAGGTAAACCTGGAATCATTCACAGACAGATTTGGAGAACTTCTTCATAAAGAT 2239
 Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGln 696
 QY 2240 GATGATGAGATGGCTCTTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
 Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
 QY 2300 GTCTGGTACAGACTTCATGACGTCTATCAACACCCCAATCTCAACAGATGATGAG 2359
 Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
 QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACACACGTGCGCAAGCCAGGACATACC 2419
 Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
 QY 2420 CCAGGAACTAAACAAATGGAAGCATTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
 Db 748 -----ProLysTrpLysHisMetGlnMetLysLysLysArgAsnArgArg 763
 QY 2480 ACCCAC---GAATTGAGAGGCGACCCAGGAGTGTCT 2512
 Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
 RESULT 12
 ABU66075 ID ABU66075 standard; Protein; 777 AA.
 XX AC ABU66075;
 XX XX
 DT 20-MAY-2003 (first entry)
 XX XX
 DE Novel human secreted and transmembrane protein PRO1491.
 KW Human; secreted protein; transmembrane protein; cytotostatic;
 KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
 KW adrenal tumour; lung tumour; colon tumour; breast tumour;
 KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
 XX Homo sapiens.
 OS
 XX US2003036157-A1.
 FN
 XX
 PD 20-FEB-2003.
 XX
 XX 02-JUL-2002; 2002US-0188769.
 XX
 XX 16-SEP-1998; 98WO-US19330.
 PR 07-OCT-1998; 98WO-US21141.
 PR 01-DEC-1998; 98WO-US25108.
 PR 08-MAR-1999; 99WO-US05028.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28551.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21086.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 29-AUG-2001; 2001WO-US27099.
 PR 18-SEP-1997; 97US-059263P.
 PR 18-SEP-1997; 97US-059266P.
 PR 17-OCT-1997; 97US-062250P.
 PR 21-OCT-1997; 97US-063486P.
 PR 24-OCT-1997; 97US-063120P.
 PR 28-OCT-1997; 97US-063121P.
 PR 28-OCT-1997; 97US-063540P.
 PR 28-OCT-1997; 97US-063541P.
 PR 28-OCT-1997; 97US-063544P.
 PR 28-OCT-1997; 97US-063564P.
 PR 29-OCT-1997; 97US-063734P.
 PR 31-OCT-1997; 97US-063870P.
 PR 31-OCT-1997; 97US-064103P.
 PR 13-NOV-1997; 97US-065311P.
 PR 21-NOV-1997; 97US-066120P.
 PR 24-NOV-1997; 97US-066466P.
 PR 11-DEC-1997; 97US-069335P.
 PR 12-DEC-1997; 97US-069425P.
 PR 17-DEC-1997; 97US-069870P.
 PR 18-DEC-1997; 97US-068017P.
 PR 10-MAR-1998; 98US-077450P.
 PR 11-MAR-1998; 98US-077632P.
 PR 11-MAR-1998; 98US-077649P.
 PR 20-MAR-1998; 98US-078886P.
 PR 20-MAR-1998; 98US-078939P.
 PR 27-MAR-1998; 98US-079664P.
 PR 27-MAR-1998; 98US-079786P.
 PR 31-MAR-1998; 98US-080107P.
 PR 01-APR-1998; 98US-080194P.
 PR 01-APR-1998; 98US-080337P.
 PR 08-APR-1998; 98US-081049P.
 PR 08-APR-1998; 98US-081070P.
 PR 09-APR-1998; 98US-081195P.
 PR 15-APR-1998; 98US-081838P.
 PR 21-APR-1998; 98US-082568P.
 PR 21-APR-1998; 98US-082569P.
 PR 22-APR-1998; 98US-082704P.
 PR 22-APR-1998; 98US-082797P.
 PR 28-APR-1998; 98US-083222P.
 PR 29-APR-1998; 98US-083495P.
 PR 29-APR-1998; 98US-083496P.
 PR 29-APR-1998; 98US-083499P.
 PR 29-APR-1998; 98US-083559P.
 PR 05-MAY-1998; 98US-084366P.
 PR 06-MAY-1998; 98US-084414P.
 PR 07-MAY-1998; 98US-084639P.
 PR 07-MAY-1998; 98US-084640P.
 PR 07-MAY-1998; 98US-084643P.
 PR 15-MAY-1998; 98US-085579P.
 PR 15-MAY-1998; 98US-085580P.

```
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087509P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088262P.
PR 05-JUN-1998; 98US-088467P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 05-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 10-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089808P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090540P.
PR 22-JUN-1998; 98US-090540P.
PR 22-JUN-1998; 98US-090540P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 26-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 04-AUG-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.

PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097552P.
PR 26-AUG-1998; 98US-097554P.
PR 26-AUG-1998; 98US-097555P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Alignment Scores:
Pred. No.: 2.68e-198
Score: 2245.50
Percent Similarity: 73.67%
Best Local Similarity: 54.92%
Query Match: 45.86%
DB: 24

US-09-774-490-1 (1-2709) x ABU66075 (1-777)

QY 275 AAGAACAAATGTCACAGGCTGAAATATCTCTACAAAGAAATGTTGGAATCCAAATGTG 334
DB 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTTGGCCACAGCTCCAGTATCATACCTTCCTTTTCGATGAGGAA 394
DB 59 IleProPheLeuGlySerSerGluGlyLeuAspGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGGAGTAGGCTGATGTTGGAGCAAGATCACAATATTTTCATTCGACCTGGTTAATATC 454
DB 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTTGTGCGCAGATCTTACACCAAGAGAGATGAATGCAAG 511
DB 99 AsnLysAsnPheLysLysIleTyrTyrProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTCGAAAGACATCTCGAAAGATGTCTAATTTTCATCAAGTACTTTAAGCATAT 571
DB 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTGTAGGCTGTGGACGGGGCTTTTCATCCAAATTTGACCTACATT 631
DB 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTCGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAAC 691
DB 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GGCGTGGGAAGATGCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
DB 179 GlyArgLeuLysCysProPheAspProGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGAACTGAGCTGATTTTATGGGCGAGACTTTGTCTATCTTCGACT 811
DB 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCAACCCCAATCAGACAGACAGCATGATTCAGGTGGCTC 862
DB 219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTTCATTAGTGCCCACTCATCTCAGAGAGTGACATCTCTGAGATGAC 922
DB 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
```

Qy	2000	TCC	CAGAGCGCGTGTCTATTGGCAATTCAGAGCGCGAAATGCAAGACGCGAAAGAAGAG	2059
Db	617	Ser	GlnGlnAlaThrIleLysTrpTrpIleGlnArgSerGlyAspGluHisArgGluGlu	636
Qy	2060	ATC	CAGAGTGATGATCATATCATCAGCAGCAGATCAAGCGCTCTCTGCTAGCTAGCTACTACAA	2119
Db	637	Leu	LysProAspGlnArgIleIleLysThrGluTrpGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAG	AAGGATTTCAGGCAATTACCTCTGCCATGCGGTGGAAACATGGGTTCATACAAACTCTT	2179
Db	657	Lys	LysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTT	AAGGTAAACCTCGAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTCATAAAGAT	2239
Db	677	Val	LysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GAT	GATGGAGATGGCTCTAAGACCAAGAATAATGTCCAATAGCATGACACCTAGCCAGAG	2299
Db	697	His	GluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTC	TGTCATCAGACAGCTTCATGAGCTCATCAACCCCAATCTCAACACGATGGATGAG	2359
Db	711	Leu	ArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTT	CTGTGAACAAGTTTGGAAAAGGGACCGAAACCAACGTGCGCAAGGCCAGGACATACC	2419
Db	730	Tyr	CysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CN	GGAACAGTAACAATGGGAAGCACTTACAAGAAATAAGAAAGGTAGAAACAGGAGG	2479
Db	748	-----	ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACC	CCAC---GAATTTGAGAGGGCACCAGGAGTGTC	2512
Db	764	His	HisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 13				
ABU67579				
XX	ABU67579	standard; Protein; 777 AA.		
XX	ABU67579;			
XX	XX	(first entry)		
XX	XX	Human secreted/transmembrane protein (PRO) #174.		
DE	XX			
XX	XX			
KW	XX	Human; secreted and transmembrane protein; PRO; TNF-alpha;		
KW	XX	tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;		
KW	XX	tissue typing.		
OS	XX	Homo sapiens.		
XX	XX			
PN	XX	US2003036162-A1.		
XX	XX			
PD	XX	20-FEB-2003.		
XX	XX			
PF	XX	12-JUL-2002; 2002US-0194423.		
XX	XX			
PR	PR	16-SEP-1998;	98WO-US19330.	
PR	PR	07-OCT-1998;	98WO-US21141.	
PR	PR	01-DEC-1998;	98WO-US25108.	
PR	PR	08-MAR-1999;	99WO-US05028.	
PR	PR	14-MAY-1999;	99WO-US10733.	
PR	PR	02-JUN-1999;	99WO-US12252.	
PR	PR	01-SEP-1999;	99WO-US20111.	
PR	PR	15-SEP-1999;	99WO-US21090.	
PR	PR	01-DEC-1999;	99WO-US28301.	
PR	PR	02-DEC-1999;	99WO-US28551.	
PR	PR	30-DEC-1999;	99WO-US31274.	
PR	PR	05-JAN-2000;	2000WO-US00219.	
PR	PR	18-FEB-2000;	2000WO-US04341.	
PR	PR	18-FEB-2000;	2000WO-US04342.	
PR	PR	22-FEB-2000;	2000WO-US04414.	
PR	PR	22-FEB-2000;	2000WO-US04414.	

Db	637	LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAGAAGATTTCAGGCAATTAACCTCTGCATCGCGTGGAAACATCGGTTCATACAAAACCTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTTGGAAAGTCATTGACACAGAGCATTTTGGAAAGAACTTCTTCATAAAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATCATGGAGATGGCTCTCTAAGACCAAAAGAAATGTCCTCAATAGCATGACACCTAGCCAGAG	2299
Db	697	HisGluGluGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGAGACTTCATCGAGCTCATCAACACCCCACTCAACACGATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTGAACAAGTTTGGAAAAGGACCGGAAACAACAGTCGCGCAAAAGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAGAAATAAGAAAGGTAGAACACGAGG	2479
Db	748	-----ProllystrLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACCAC---CAATTTGAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 14			
ABU65437	ID	ABU65437 standard; Protein; 777 AA.	
XX	AC	ABU65437;	
XX	DT	16-MAY-2003 (first entry)	
XX	DE	Human PRO polypeptide #174.	
XX	KW	Human; PRO; cytostatic; chromosome mapping; gene mapping;	
KW	KW	protein electrophoresis; tumour necrosis factor-alpha; TNF-alpha; blood;	
KW	KW	chondrocyte differentiation; chondrocyte proliferation; tumour.	
XX	OS	homo sapiens.	
XX	PN	US2003032102-A1.	
XX	PD	13-FEB-2003.	
XX	PF	17-JUN-2002; 2002US-0173697.	
XX	XX	16-SEP-1998; 98WO-US19330.	
PR	PR	07-OCT-1998; 98WO-US21141.	
PR	PR	01-DEC-1998; 98WO-US25108.	
PR	PR	08-MAR-1999; 99WO-US05028.	
PR	PR	14-MAY-1999; 99WO-US10733.	
PR	PR	02-JUN-1999; 99WO-US12252.	
PR	PR	01-SEP-1999; 99WO-US20111.	
PR	PR	15-SEP-1999; 99WO-US21090.	
PR	PR	01-DEC-1999; 99WO-US28301.	
PR	PR	02-DEC-1999; 99WO-US28551.	
PR	PR	30-DEC-1999; 99WO-US31274.	
PR	PR	05-JAN-2000; 2000WO-US00219.	
PR	PR	18-FEB-2000; 2000WO-US04341.	
PR	PR	18-FEB-2000; 2000WO-US04342.	
PR	PR	22-FEB-2000; 2000WO-US04414.	
PR	PR	24-FEB-2000; 2000WO-US05004.	
PR	PR	01-MAR-2000; 2000WO-US05601.	
PR	PR	02-MAR-2000; 2000WO-US05841.	
PR	PR	13-MAR-2000; 2000WO-US06884.	
PR	PR	30-MAR-2000; 2000WO-US08439.	
PR	PR	17-MAY-2000; 2000WO-US13705.	

PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US05520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059466P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 09-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 02-JUL-1998; 98US-091633P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.

279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuLeuAsn	298
1043	AAATGGACACATCTCTCAAGTCCTCTGATTTGCTCAGTGCAGGTCCAAATGCGATT	1102
299	LysTrpThrPheLeuLysAlaArgLeuLeuCysSerIlePheGlySerAspGlyAla	318
1103	GACACTCATTTTCATCAACTGCAGGATGTATCTCTAATGAACCTTAAAGATCCTAAAAAT	1162
319	AspThrTyrrPheAspGluLeuGlnAspIleTyrrLeuLeuProThrArgAspGluArgAsn	338
1163	CCAGTTGTATATCGAGTGTTTACGACTCCAGTAAACATTTTCAAGGATCAGCCGTGTGT	1222
339	ProValValTyrrGlyValPheThrThrThrSerSerIlePheLysGlySerAlaValCys	358
1223	ATGTATAGCATGATGTGAGAAGGGTGTCTTGTGTCATATCCACACGGGATGGA	1282
359	ValTyrrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrrAlaHisLysGluSer	378
1283	CCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCCTATCCAGCGCCAGCAACTGT	1342
379	AlaAspHisArgTrpValGlnTyrrAspGlyArgIleProTyrrProArgProGlyThrCys	398
1343	CCGACGAAAACATTTGGTGGT---TTTGACTCTCAAAAGAGACCTTCTGATGATGTATA	1399
399	ProSerLysThrTyrrAspProLeuLysSerThrArgAspPheProAspAspValIle	418
1400	ACCTTTTGCAAGAAGTATCCAGCCATGTACATCCAGTGTTCCTATGAACAATCGCCCA	1459
419	SerPheIleLysArgHisSerValMetTyrrLysSerValTyrrProValAlaGlyGlyPro	438
1460	ATAGTGATCAAAACGGATGTAAATATCAATTATACACAAATTCCTGTAGACCCAGTGGAT	1519
439	ThrPheLysArgIleAsnValAspTyrrArgLeuThrGlnIleValValAspHisValIle	458
1520	GCAGAAGATGGACAGTATGATGTATGTTTATCGGAACAGATGTGGGACCGTCTTTAAA	1579
459	AlaGluAspGlyGlnTyrrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys	478
1580	GTAATTTCAATTCCTTAAGCAGACTTGTATGATTTAGAGAAGGTCTCTGTGGAGAAGATG	1539
479	ValValSerIleSerLysGluLysStrp---AsnMetGluGluValValLeuGluGluLeu	497
1640	ACAGTTTTTCGGAAACCGACTGCTATTTCAGCAATCGAGCTTTCACCTAAGCAGCAACAA	1699
498	GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln	517
1700	CTATATATTGGTTTCAACGGCTGGGTGGCCAGCTCCCTTTTACACCGGTGTGATATTTAC	1759
518	LeuTyrrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyrr	537
1760	GGGAAACGTGTGCTCAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGATGTTCT	1819
538	GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrrCysAlaTrpAspGlyAsn	557
1820	GCATGTTCTCGCTATTTCCTCCACTGCAAGAGACGCACAAAGACGACAAGATATAAGAAAT	1879
558	AlaCysSerArgTyrrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyrr	577
1880	GGAGACCCACTGACTCATTCTGTTTCAGACTTTACACCATGATAATCCACCTGCCACAGCCCT	1939
578	GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla	596
1940	GAAGAGAGAATCATCTATGCTGTAGAGAAATAGTAGCACATTTTGGAAATGCGAGTCCGAAG	1999
597	AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys	616
2000	TGCGACAGAGCGCTGTCTATTGGCAATTCACAGACGCGAAATGAAGACGCGAAAGAAGAG	2059
617	SerGlnGlnAlaThrIleLysTrpTyrrIleGlnArgSerGlyAspGluHisArgGluGlu	636
2060	ATCAGAGTGCATCATCATCATCAGGACAGATCAAGGCTTCTGTCTAGTGTCTACAA	2119

Db	637	LeuLysProAspGluArgLlelleLysThrGluTyrGlyLeuLeulleArgSerLeuGln	656
Qy	2120	CAGAAGGATTGAGCAATTAACCTCTGCATCGCTGGAAACATGGTTTCATACAAACCTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGTAAACCTCGAAGTCATTGACACAGACGATTGGGAAGAACTTCTTCATAAAGAT	2239
Db	677	VallLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGAGACTTCATCGACTCATCAACCAACCCCAATCTCAACACGATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTCAACAAGTTTGGAAAGGACCGAAAAACAACGTCGGCAAGGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTyrPheArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGACACAGTACAAATGGAAGCACTTACAGAAAAATAGAAAAGGTAGAAACAGGAGG	2479
Db	748	-----ProLysTyrLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACCCAC--GAATTTGAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 15			
ABUS8573			
ID	ABUS8573 standard; Protein; 777 AA.		
AC	ABUS8573;		
DT	15-APR-2003 (first entry)		
XX	Human PRO polypeptide #174.		
DE	Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach;		
KW	liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;		
KW	antibody-dependent enzyme mediated prodrug therapy.		
OS	Homo sapiens.		
XX	US2003027272-A1.		
XX	06-FEB-2003.		
XX	21-JUN-2002; 2002US-0176492.		
XX	16-SEP-1998; 98WO-US19330.		
PR	07-OCT-1998; 98WO-US21141.		
PR	01-DEC-1998; 98WO-US25108.		
PR	08-MAR-1999; 99WO-US05028.		
PR	10-MAR-1999; 99WO-US05190.		
PR	14-MAY-1999; 99WO-US10733.		
PR	02-JUN-1999; 99WO-US12252.		
PR	01-SEP-1999; 99WO-US20111.		
PR	15-SEP-1999; 99WO-US21090.		
PR	30-NOV-1999; 99WO-US28313.		
PR	01-DEC-1999; 99WO-US28301.		
PR	02-DEC-1999; 99WO-US28551.		
PR	30-DEC-1999; 99WO-US31274.		
PR	05-JAN-2000; 2000WO-US00219.		
PR	18-FEB-2000; 2000WO-US04341.		
PR	18-FEB-2000; 2000WO-US04342.		
PR	22-FEB-2000; 2000WO-US04414.		
PR	24-FEB-2000; 2000WO-US05004.		
PR	01-MAR-2000; 2000WO-US05841.		
PR	02-MAR-2000; 2000WO-US05601.		
PR	10-MAR-2000; 2000WO-US06319.		


```
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuLeuLeu 298
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1043 AAATGGACAACTTCCTCAAGCTCCTGATTTGCTCAGTCCAGGTCCAAATGCGATT 1102
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 299 LysTrpThrPheLeuLysAlaArgLeuLeuCysSerIleProGlySerAspGlyAla 318
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1103 GACACTCATTTGATGAACCTGCGAGATGATTCCTTAATCAACTTTAAGATCCTAAAT 1162
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1163 CCAAGTTGATATGAGTGTTCAGACTTCAGTAACATTTTCAAGGATCAGCGGTGT 1222
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1223 ATGTATAGCATGATGATGAGAGGGTTCCTTTGGTCCATATGCCACAGGATGA 1282
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1283 CCCAACTATCAATGGTGGCTTATCAAGGAGAGTCCCTATCCACGGCCAGGAATTGT 1342
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1343 CCAGCAAAACATTTTGGTGT---TTTGACTCTCAAAAGGACCTTCTCATGATGTATA 1399
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 399 ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle 418
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1400 ACTTTGCAAGATGATCCAGCCATGATCAATCCAGTGTTCCTATGAACAAATCCGCA 1459
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1460 ATAGTATCAAAACGATGATTAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1520 GCAGAGATGGACAGTATGATGTATGTTTATCGGAACAGATGTTGGGACCGTCTTAA 1579
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1580 GTATTTCAATTCCTTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCTGGAAGAAATG 1639
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluLeu 497
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAA 1699
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 498 GlnIlePheLysHisSerSerIleLeuLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1700 CTATATATTGGTTCAACGGTGGGGTTCGCCAGCTCCCTTTACACGGGTGATATTAC 1759
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1760 GGAAGAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTTGGATGCTTCT 1819
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACCGCACAGACAGCAAGATATAGAAAT 1879
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1880 GGAGACCCACTGACTCACTGTTCCAGATTTACACCATGATATACCATGGCCACAGCCCT 1939
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1940 GAAGAGAGATCATCTATGCTGATAGATAATAGTACACATTTTGAATGCGAGTCCGAG 1999
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2000 TCCAGAGAGCGTGTCTTATGGCAATTCAGAGCGGAAATGAGAGCGAAAGAGAG 2059
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2060 ATCAGAGTGGATCATCATCATCAGGACAGATCAAGCGCTTCTGCTAGTGTCTACAA 2119
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

Search completed: August 3, 2003, 10:05:38
Job time : 127.5 secs

```
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAGAGATTTCAGCAATTACTCTGCCATCGGTGGACATGGTTTCATACAACTCTT 2179
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2180 CTTAAGGTAAACCTGGGAAGTCAATGACACAGAGCATTTGGGAAGAACTTCTTCATAAAGAT 2239
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2240 GATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2300 GTCTGGTACAGAGACTTCATCAGTCAATCAACACCCCAATCTCAACACAGTGGATGAG 2359
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2360 TTCTGTGAACAAAGTTTGGAAAAGGACCGGAAACAAACGTCGGCAAGGCCAGGACATACC 2419
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2420 CCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGAGG 2479
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2480 ACCCAC---GAATTTGAGAGGCGCCACCCAGGAGTGTCT 2512
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search time 51 Seconds
(without alignments)
10216.508 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttcatgatg.....aggcttttttctcctaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2.1/USPTO/spool/US09774490/runat 03082003 095416 25340/app_query.fasta_1.2887
-DB=PIR 76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1.1.69 @runat 03082003 095416 25340 -NCPU=6 -ICPU=3
-NO.MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: PIR 76:*
2: PIR 1:*
3: PIR 2:*
4: PIR 3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4201	85.8	771	2	D49423
2	3989.5	81.5	772	2	I48747
3	3784.5	77.3	772	2	A49069
4	3478.5	71.0	666	2	I58169
5	2118	43.3	749	2	G01856
6	2059.5	42.1	748	2	I48744
7	1952	39.9	753	2	G02173
8	1850.5	37.8	751	2	I48748
9	966	19.7	834	2	S66498
10	857.5	17.5	782	2	I48746
11	802	16.4	760	2	I48745
12	785.5	16.0	724	2	C49423
13	766.5	15.7	730	2	JH0798
14	731.5	14.9	656	2	B49423

15	730	14.9	1074	2	JC5928	semaphorin F precu
16	699.5	14.3	712	2	T27165	hypothetical prote
17	684.5	14.0	711	2	A49423	semaphorin I precu
18	534	10.9	653	2	T03102	semaphorin homolog
19	327	6.7	676	2	T33853	hypothetical prote
20	197	4.0	1894	2	JC4980	plexin 1 precursor
21	188.502	3.9	1945	2	T13937	plexin A - fruit f
22	186	3.8	1872	2	JC4976	plexin 3 precursor
23	184	3.8	441	2	S29921	hypothetical prote
24	180	3.7	403	2	E42521	A39R protein - vac
25	172	3.5	1884	2	JC4975	plexin 2 precursor
26	170	3.5	1905	2	T51553	Plexin - African c
27	167.002	3.4	2051	2	T13164	plexin B - fruit f
28	154.5	3.2	1568	2	T09074	semaphorin recepto
29	134.5	2.7	1806	2	T23298	hypothetical prote
30	127.5	2.6	295	2	J01775	Salt9R protein - v
31	123	2.5	573	1	CSBYT	catalase (EC 1.11
32	120	2.5	904	2	S53856	DNA mismatch repai
33	119	2.4	1104	1	A36866	microbial collagen
34	118.5	2.4	1291	2	T09273	probable tail-host
35	118.5	2.4	1654	2	A12067	two-component sens
36	118	2.4	1379	1	S01254	hepatocyte growth
37	117.5	2.4	1029	2	H86179	hypothetical prote
38	115	2.3	3381	2	T42389	versican precursor
39	114.502	2.3	993	2	A36873	protein-tyrosine k
40	114	2.3	692	1	S57592	probable phosphoes
41	113	2.3	1276	2	T09204	probable tail-host
42	112.5	2.3	248	2	A46652	glucosamine-6-phos
43	112	2.3	4273	2	C59679	polyketide synthas
44	111.5	2.3	6642	2	T29757	protein UNC-89 - C
45	111	2.3	1375	2	T30813	plasmaonogen relate

ALIGNMENTS

RESULT 1

D49423

semaphorin III precursor - human

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999

C:Accession: D49423

R:Kotodkin, A.D.; Matthes, D.J.; Goodman, C.S.

Cell 75, 1389-1399, 1993

A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone

A:Reference number: A49423; MUID:94094332; PMID:8269517

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-771 <KOL>

A:Cross-references: GB:L26081; NID:G799328; PIDN:AAA65938.1; PID:G436560

C:Genetics:

A:Gene: GDB:SEMA1

A:Cross-references: GDB:283448

C:Superfamily: semaphorin

Alignment Scores:

Pred. No.: 1.59e-310 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D49423 (1-771)

QY 200 ATGGCGTGGTAACTAGAGTGTCTTTCTGGGAGTATTACTTACAGCAAGACA 259

|||||

Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGATGGAGACAAATGTCACAGGCTGAATTTATCTCTACAAAGAATGTTG 319

|||||

Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

320 GAATCCAAATGATGATCACTTTCAATGGCTGGCCAAACAGCTCCAGTTATCATCTTC 379
Db |||||ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTTCGATGAGGAACGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Qy |||||ATAGTGATCAAAACGATGTAATATCAATTTACAAATTTGTCGTAGACCGAGTGAT 1519
Db |||||IleValIleThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaAlaAspHisIlePheSerPhe 80
Qy |||||GCAGNAGTGCACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTAAA 1579
440 GACCTGGTAAATATCAAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACGAGA 499
Db |||||AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
81 AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArgArg 100
Qy |||||GTAGTTTCAATTCCTAAGGAGACTTGGTATCATTTAGAAGAGTTCTCTGTGAAGAAATG 1639
Db |||||ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluLysMet 480
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
Qy |||||ACAGTTTTTCCGGAAACCGACTGCTATTTACAGCAATGGAGCTTTCACTAAGCAGCA 1699
Db |||||ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
Qy |||||CTATATATTGTTTCAACGGCTGGGTTGCCAGCTCCCTTACACCGGTGTGATATTAC 1759
Db |||||LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
620 TGCACCTACATTGAATATGACATCATCTCTGAGGACAATATTTTAAAGCTGAGAACTCA 679
Qy |||||GGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTTGGGATGGTCT 1819
Db |||||CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
680 CATTTTGAACGGCGTGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
Qy |||||HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
Qy |||||TTAATAGATGGAATATATCTCTGGAACCTGACGCTGATTTTATGGGCGAGACTTGTCT 799
Db |||||LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTCCGAACTCTTGGGACACACCAATCAGGACAGAGAGATGATTCAGAGTGG 859
Qy |||||IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
860 CTCATGATGCCAAGTTCTATTAGTGCACCTCATCTCAGAGAGTGACATCTCCTGAAGAT 919
Db |||||LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
920 GACAAAGTATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTGGAAAGCT 979
Db |||||AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGATAGTGCAGATATCCAAAGATGACTTTGGAGGGGACAGAGTCTGGTG 1039
Qy |||||ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAATGACACATCTCTCAAAGCTCGTCTGATTTGCTCAGTGCAGGTCCCAATGGC 1099
Qy |||||AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTCATGAACCTGACAGGATGTATCTCTAATGAACCTTTAAAGATCCTAA 1159
Db |||||301IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 ATTCAGTGTATATGAGTGTATTCAGACTTCAGTAAACATTTTCAAGGGATCAGCCGNG 1219
Qy |||||321AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGTATAGCATGATGTGAGAAGGGTGTCTTGGTCCATATCTCCATATCCACAGGGAT 1279
Db |||||341CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGGCAGGAAT 1339
Qy |||||361GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValPheLeuGlyProTyrAlaHisArgAsp 380
1340 TGTCCAGCAAAACATTTGGTGGTGTGATCTACAAAGGACCTCTCTGATGATGTTATA 1399
Qy |||||381CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
Qy |||||1400ACCTTTGCAAGAAGTATCCAGCCATGTACAATCCAGTGTCTTCTATGAACAATCGCCCA 1459

401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy |||||1460ATAGTGATCAAAACGATGTAATATCAATTTACAAATTTGTCGTAGACCGAGTGAT 1519
Db |||||421IleValIleThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy |||||1520GCAGNAGTGCACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTAAA 1579
Db |||||441AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
Qy |||||1580GTAGTTTCAATTCCTAAGGAGACTTGGTATCATTTAGAAGAGTTCTCTGTGAAGAAATG 1639
Db |||||461ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluLysMet 480
Qy |||||1640ACAGTTTTTCCGGAAACCGACTGCTATTTACAGCAATGGAGCTTTCACTAAGCAGCA 1699
Db |||||481ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
Qy |||||1700CTATATATTGTTTCAACGGCTGGGTTGCCAGCTCCCTTACACCGGTGTGATATTAC 1759
Db |||||501LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
Qy |||||1760GGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTTGGGATGGTCT 1819
Db |||||521GlyLysAlaCysAlaGluCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
Qy |||||1820GCATGTTCTCCTATTTTCCACTGCAAGAGACGACACAGACGACAGATATAAGAAAT 1879
Db |||||541AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGluAspIleArgAsn 560
Qy |||||1880CGAGACCCACTGACTCTGCTGACTTACACCATGATATCACCATGGCCACACCCCT 1939
Db |||||561GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
Qy |||||1940GAAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTTCGAAATGAGTCCGAG 1999
Db |||||581GluGluAlaGlyIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
Qy |||||2000TCGACAGAGCGCTGCTCTATTGGCAATTCAGAGCGGCAATGAAGAGCGGAAAGAG 2059
Db |||||601SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
Qy |||||2060ATCAGACTGAGTATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAA 2119
Db |||||621IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy |||||2120CAGAAGGATTCAGGCAATTTACTCTCTGCCATCGGTGGAACATGGGTTTCATACAAACTCTT 2179
Db |||||641GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
Qy |||||2180CTTAAGTAACTCCGGAAGTCAATGACACAGACATTTGGAGAACTTCTTCATAAAGAT 2239
Db |||||661LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
Qy |||||2240GATGATGAGATGGCTCTTAAGACCAAGAAATGCCATAGCATGACACCTAGCCAGAG 2299
Db |||||681AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
Qy |||||2300GTCTGGTACAGAGACTTCATGCAGCTCATCAACACCCCAATCTCAACACCATGGATGAG 2359
Db |||||701ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
Qy |||||2360TTCGTGTAACTGAAAGGACCCGAAACAACTCGGCAAGGCCAGGACATACC 2419
Db |||||721PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
Qy |||||2420CCAGGACACTTAACTGAGGACCTTACAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
Db |||||741ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
Qy |||||2480ACCCACGAATTTGAGAGGGCCACCCAGGAGTGTCT 2512

Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
I48747
semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48747
R:Fuscher, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48747
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-772 <RES>
A:Cross-references: EMBL:X85993; NID:g854329; PIDN:CAA59985.1; PID:g854330
C:Genetics:
A:Gene: smd
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 1,76e-294 Length: 772
Score: 3989.50 Matches: 730
Percent Similarity: 97.28% Conservative: 21
Best Local Similarity: 94.56% Mismatches: 20
Query Match: 81.48% Indels: 1
Gaps: 2

US-09-774-490-1 (1-2709) x I48747 (1-772)

QY 200 ATGGCTGGTTAACTAGGATTGTCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGACAACTGTGCCAGGCTGAATATCTCTCAAGAAGATGTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAACTATGTCATCTTCAATGGCTGGCCCAACAGCTCCAGTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTGGCAAGAGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTATATCAAGATTTTCAAAGATTGTGTGCCAGTATCTTACACAGAAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCGAAGAAATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelIleVal 120
QY 560 CTTAAGCATATAATCAGACTCACCTGTACGCTGTGGAACGGGGCTTTTCATCCAAT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAATTTGGACATCATCTCGAGGACAATATTTTAACTGGAGACTCA 679
Db 141 CysThrTyrIleGluValGlyHisHisProGluAspAsnIlePheLysLeuGlnAspSer 160
QY 680 CATTTTGAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTCACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAGATTAATFACTCTCGAACTGCAGCTGATTTTATGGGCGAGACTTCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTTGGGCACCAACCAATCAGGACAGACAGCATGATTCACAGTGG 859

Db 201 IlePheArgThrLeuGlyAspHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCCTGAAGAT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
QY 920 GACAAAGATATCTTTTCTCCGTGAAATCAATAGATGAGAACACTCTGGAAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleGlyGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTGAGATAGTTCAGATATGCAAGATGCTTGGAGGCGCAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAGCTCGTCTGATTTTCTCAGTCCAGGTCCTCAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTAGTAACTGCAGATGTATTCCTTAATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTTCAGACTTCAGTAACTATTCAGGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGCTGTCTTGTGCTCATATGCCACACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCGCAAAACATTTGGTGGTTTGTACTCTCAAGAGGACCTCTCTGATGCTTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGGCAAGACTCATCCAGCCATGTACAACTCCAGTGTTCCTATCAACAAATCGCCA 1459
Db 401 ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAAATTTATCAATTTACAAATTTGCTGAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGACAGTATGATGTTATGTTATCGAACAGATGTGTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGTCTATTTCAGCAATGGAGCTTTCCTCAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTTGGTTCACCGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGTGCTCGCCGAGACCTTACTGTGCTGGATGGTGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCCAAGAGACGCAAGACGACAGATATAGAAT 1879
Db 541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTACTGTTTCAAGCTTACAC---CATGATATCATCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisHisGlyProSer 580

```
QY 1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCGTCCG 1996
Db 581 LeuGluGluArgGlyLeuValGluAsnSerThrPheLeuGluCysSerPro 600
QY 1997 AAGTCCAGAGAGCGCTGCTATTGCGCAATCCAGAGCGCAAAATGAAGAGCGAAAGAA 2056
Db 601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnArgSerLysArg 620
QY 2057 GAGATCAGAGTATGATCATATCATCATCAGACAGATCAAGCGCTCTCTGTAGTGTCTA 2116
Db 621 GluIleArgMetGlyAspHisIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
QY 2117 CACAGAGAGTATCAGCAATACCTCTCCATGCGGTGGAGACATGGTTCATACAACT 2176
Db 641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
QY 2177 CTTCTTAAGCTAACCTCGAAGTTCATTGACACAGACATTTGGAAGAACTTCTTCATAAA 2236
Db 661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLys 680
QY 2237 GATGATGATGAGATGGCTCTTAAGACCAAGAAATGCTCAATAGATGACACCTAGCCAG 2296
Db 681 AspAspGlyAspGlySerLysIleLysGluMetSerSerMetThrProSerGln 700
QY 2297 AAGGTCTGTACAGACTTCATGAGCTCATCAACCCCACTCAACAGCATGGAT 2356
Db 701 LysValTyrPyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY 2357 GAGTCTGTGAACAACTTTGGAAAGGACCGCAACACAGCTCGGCAAGGCGCAGACAT 2416
Db 721 GluPheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHis 740
QY 2417 ACCCAGGAGACAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAACAGG 2476
Db 741 SerGlnGlySerSerAsnLysTyrLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY 2477 AGGACCCAGCAATTTGAGAGGCGCACCCAGGAGTGC 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 3
A49069
collapsin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
C:Accession: A49069
R:Ruo, Y.; Raible, D.; Raper, J.A.
Cell 75, 217-227, 1993
A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neurons
A:Reference number: A49069; MUID:94006554; PMID:8402508
A:Accession: A49069
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-772 <LUO>
A:Cross-references: GB:U02528; NID:G410078; PIDN:AAC59638.1; PID:G410079
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 6,29e-279 Length: 772
Score: 3784.50 Matches: 683
Percent Similarity: 95.34% Conservative: 53
Best Local Similarity: 88.47% Mismatches: 35
Query Match: 77.30% Indels: 1
DB: 2 Gaps: 1

US-09-774-490-1 (1-2709) x A49069 (1-772)

QY 200 ATGGGCTGCTTAACCTAGGATGTCTGCTTTCTGGGAGTATTACTTACACAGAGCA 259
Db 1 MetGlyTyrLeuArgGlyIleAlaLeuLeuSerLeuGlyValLeuLeuAlaGlyArgVal 20
QY 260 AACTATCAGATGGGAGAACAACTGTGCCAAGCTGAAATTTCTCTACAAAGAAATGTTG 319
```

```
Db 21 AsnCysGlnHisValLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACATGTGATCATCTTTCAATGGCTTGGCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnIleValAsnPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAGCGAGTAGGCTGTATGTCGAGCAAAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 CACTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAAGA 499
Db 81 AsnLeuValAsnIleLysGluTyrGlnLysIleValTyrProValSerHisSerArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGCTTAATTTTCATCAAGTGA 559
Db 101 AspGluCysLysTyrAlaGlyLysAspIleLeuArgGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATCTGTACCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysThrTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProMet 140
QY 620 TGCACCTACATTTGAAATTTGGACATCTCTCAGGACCAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluValGlySerHisProGluAspAsnIlePheArgMetGluAspSer 160
QY 680 CATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGATATATCTCTGGAACTCAGCTGATTTTATGGGCGGAGACTTTGCT 799
Db 181 LeuValAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGCCACCCACCAATCAGGACAGACAGCATGATTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisPheIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCCGAAAGAT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTCCGTAATAATCCAATAGATGAGAGAACACTCTGGAAGAGCT 979
Db 241 AspLysIleTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisThrGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATATGCAAGAAATGACTTTGGAGGCGCACAGAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACTTCTCAAAAGTCTGCTGATTGCTCAGTCCAGGTCCCAATGCG 1099
Db 281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGATGAAGTCAAGTATGTTTCTTAATGAACCTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTATTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATAGCATGATGATGTGAGAAGGTGTCTTGGTCCATATGATCCACAGGAT 1279
Db 341 CysMetTyrSerMetThrAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTTATCAAGGAAGAGTCCCTATATCCACGGCCAGCAACT 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACCTTCTGTGATGATTTATA 1399
```

Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspGluValIle 400
QY 1400 ACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyraenProValPheProIleasnSerArgPro 420
QY 1460 ATAGTGATCAAAACCGATGAATATATCAATTTACAAATTTGCTGAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAspTyrglnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAAGATGACAGTATGATGATTTATCGGAACAGATGTTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyraaspValmetPheIleGlyThrAspIleGlyThrValIleuLys 460
QY 1580 GPAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGlnThrTrpHisGluLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTCTATTTACGCAATGGAGCTTTCACCTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrValIleSerAlaMetLysIleSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGGTTCACCGCTGGGTTCGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuIleGlySerAlaThrGlyValSerGlnLeuProLeuHisArgCysAspValTy 520
QY 1760 GCGAAGCGGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyraaspIlePheGlySer 540
QY 1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACGCCAAGACGACACAGATATAGAAT 1879
Db 541 SerCysSerArgTyraPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTTACACTTA---CACCATGATTAATCACCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGlnHisAspAsnProSerGlyGlnThr 580
QY 1937 CCTGAAGAGAGAAATCATCTATGTTGTAGAGATAGTAGACATTTTTCGAATTCAGTCCG 1996
Db 581 LeuGluGluLysIleIleTyraGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
QY 1997 AAGTCGAGAGAGCGCTGTATTTGGCAATTCAGAGCGCAATGAAGAGCGCAAAAGAA 2056
Db 601 LysSerGlnArgAlaIleValTyraPheGlnPheGlnLysGlnAsnAspAspHisLysVal 620
QY 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTCTACGTAGTCTA 2116
Db 621 GluIleLysValAspAspArgMetIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
QY 2117 CAACAGAAGGATTCAGGCAATTTACCTCTGCCATGGGTGGAAACATGGGTTCATACAAACT 2176
Db 641 GlnArgArgAspSerGlyIleTyraPheCysHisAlaValGluHisGlyPheIleGlnThr 660
QY 2177 CTTCTTAAGGTAAACCTCGAAGTCAATGACACAGAGCATTTTGAAGAACTTCTTCAATAA 2236
Db 661 LeuLeuLysValThrLeuGluValIleAspThrAspHisLeuLeuGluLeuLeuHisLys 680
QY 2237 GATGATGATGAGATGGTCTTAAGACCAAGAAATGTCATAGATGACACCTAGCCAG 2296
Db 681 GluGluAspAlaAspLysThrLysAspAlaThrAsnSerMetThrProSerGln 700
QY 2297 AAGTCTCGTACAGAGACTTCATGAGCTTCATCAACACCCCACTCTCAACACCATGAT 2356
Db 701 LysIleTrpTyraaspPheMetGlnIleuIleasnHisProasnLeuAsnThrMetAsp 720
QY 2357 GAGTTCCTGCAACAGTTTGGAAAAGGACCCGAAACACCTCGCGCAAGGCCCGAGCAT 2416
Db 721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProAlaasn 740
QY 2417 ACCCCAGGGAACAGTAACAATGGAGACCTTACAGAAATATAGAAAGGTAGAAACAGG 2476
Db 741 AlaGlnValAsnThrAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArg 760

QY 2477 AGGACCCACGAATTTGAGAGGGCACCCAGGAGTGTG 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 4

158169 semaphorin III - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999

C:Accession: 158169

R:Mesersmith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; Ko

Neuron 14, 949-959, 1995

A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory pr

A:Reference number: 158169; MUID:95267432; PMID:7748562

A:Accession: 158169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-666 <RES>

A:Cross-references: GB:L40484; NID:g703189; PIDN:AAA73934.1; PID:g703190

C:Genetics:

A:Gene: SemaIII

C:Superfamily: semaphorin

Alignment Scores:

Pred. No.:	1,01e-255	Length:	666
Score:	3478.50	Matches:	633
Percent Similarity:	98.50%	Conservative:	23
Best Local Similarity:	95.05%	Mismatches:	9
Query Match:	71.05%	Indels:	1
DB:	2	Gaps:	1

US-09-774-490-1 (1-2709) x 158169 (1-666)

QY 518 GGAAGAAGACATCTGGAAGAATGTCTAATTTTCATCAAGGTACTTAAGGCATATATCAG 577
Db 1 GlyLysAspIleLeuLysGluCysAlaAsnPheIleLysValLeuGluAlaTyraAsnGln 20
QY 578 ACTCACTGTGACGCTGTGGAACCGGGGCTTTTCATCAATTCGACCTACATGCAAAAT 637
Db 21 ThrHisLeuTyraLysGlyThrGlyAlaPheHisProIleCysThrTyraIleGluVal 40
QY 638 GGACATCATCTGAGGACAATATTTTAAGCTGAGAACTCACATTTTGAAGACGGCGCT 697
Db 41 GlyHisProGluAspAsnIlePheLysLeuGlnAspSerHisPheGluAsnGlyArg 60
QY 698 GGGAGAGTCCATATGACCCCTAAGCTCTGACACATCCCTTTTAATAGATGGAATA 757
Db 61 GlyLysSerProTyraAspProLysLeuLeuThrAlaSerLeuLeuIleAspGlyGluLeu 80
QY 758 TACTCTGAACTGACGCTGATTTTATGGGCGAGACTTTTGTATCTTCCGAACCTCTTGGG 817
Db 81 TyraSerGlyThrAlaAlaAsnPheMetGlyArgAspPheAlaIlePheArgThrLeuGly 100
QY 818 CACCACCAACCAATCAGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTC 877
Db 101 HisHisPheProIleArgThrGluGlnHisAspSerArgTrpLeuAsnAspProArgPhe 120
QY 878 ATTAGTCCCACTTCATCTCAGAGAGTGCACATCTCTGAAGATGACAAAGTATATCTTTTC 937
Db 121 IleSerAlaHisLeuIleProGluSerAspAsnProGluAspAspLysValTyraPhePhe 140
QY 938 TTCGCTGAAATCAATAGATGGAACACCTCTCGAAAGCTACTCTACGCTAGATAGGT 997
Db 141 PheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAlaThrHisAlaArgIleGly 160
QY 998 CAGATATCAAGATGACTTTTGGAGGGCACAGAGTCTGGTGAATAATAATGACACATTC 1057
Db 161 GlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuValAsnLysTrpThrPhe 180
QY 1058 CTCAAGCTCGTGTGATTTGCTCAGTCCAGGTCCAAATGGCATTTGACATCATTTGAT 1117
Db 181 LeuLysAlaArgLeuIleCysSerValProGlyProAsnGlyIleAspThrHisPheAsp 200

A:Reference number: I48744; MUID:95267431; PMID:7748561
 A:Accession: I48744
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-748 <RES>
 A:Cross-references: EMBL:X85990; NID:g854323; PIDN:CAAS9982.1; PID:g854324
 C:Genetics:
 A:Gene: sema
 C:Superfamily: semaphorin

Alignment Scores:
 Pred. No.: 4,62e-148 Length: 748
 Score: 2059.50 Matches: 400
 Percent Similarity: 66.93% Conservative: 112
 Best Local Similarity: 52.23% Mismatches: 184
 Query Match: 42.06% Indels: 69
 DB: 2 Gaps: 15

US-09-774-490-1 (1-2709) x I48744 (1-748)

QY	281	ANTGTCGAAGCTGAATATTCCTACAAAGAAATGTTGGAAATCCAAATGTGATCACT	340	
DB	27	AsnLeuProArgLeuArgLeuSerPheGlnGluLeuGlnAlaArgHisGlyValArgThr	46	
QY	341	TTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGAACGGAGT	400	
DB	47	PheArg---LeuGluArgThrCysCysTyrGluAlaLeuLeuValAspGluArgGly	65	
QY	401	AGGCTGTATGTTGGCAAGGATCACATATTTTCATTCGACCTGGTTAATATC---AAG	457	
DB	66	ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuSerLeuAspAsnIleSerLys	85	
QY	458	GATTTTCAAAAGATGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCT	517	
DB	86	ArgAlaLysLysLeuAlaTrpAlaProValGluTrpArgGluGluCysAsnTrpAla	105	
QY	518	GGAAAGACATCTGAAGAAGTGTGCTAATTTTCATCAAGGTACTTAAGCATATATCAG	577	
DB	106	GlyLysAspIleGlyThrGluCysMetAsnPheValArgLeuLeuHisAlaTyrAsnHis	125	
QY	578	ACTCACTGTACGCTGTGGAAAGCGGGGCTTTTCATCCAAATTCG-----	622	
DB	126	ThrHisLeuLeuAlaCysArgThrGlyAlaPheHisProThrCysAlaLeuTrpArgTrp	145	
QY	623	-----ACCTACATTGAAATGGACATCATCTCTGAGACAAATTTTAAG	667	
DB	146	AlaThrAlaGlyGlyThrHisAlaSerThrGly-----ProGlu-----	158	
QY	668	CTGGAGAACTCACATTTTCAAAAGCGCGTGGAGAGTCCATATGACCTTAAGCTGCTG	727	
DB	159	-----LysLeuGluAspGlyLysGlyLysThrProTyrAspProArgHisArg	174	
QY	728	ACAGCATCCCTTTTAATAGATGAGAAATATATCTCTGGAACCTGCAGCTGATTTATGGGG	787	
DB	175	ProProSerValLeuValGlyGluGluLeuTyrSerGlyValThrAlaAspLeuMetGly	194	
QY	788	CGAGACTTGTCTATCTTCGAACTCTTGGGACCCACCACCAATACGACAGAGCAGAT	847	
DB	195	ArgAspPheThrIlePheArgSerLeuGlyGlnAsnProSerLeuArgThrGluProHis	214	
QY	848	GATTCAGGTGGCTCAATCATCAAAAGTTTCATTAGTGGCCCACTCATCTCAGAGATGAC	907	
DB	215	AspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSerGlu	234	
QY	908	AATCTCGAAGATGACAAAGTATATCTTTCTTCGTTGAAAATGCAATGAGGAAACAC	967	
DB	235	AsnProAspAspAspLysIleTyrPhePhePheArgGluSerAlaValGluAlaPro	254	
QY	968	TCT---GGAAAGCTACTCAGCTAGAAATAGTTCAGATATGCAAGATCACTTTGAGGG	1024	
DB	255	AlaMetGlyArgMetSerValSerArgValGlyGlnIleCysArgAsnAspLeuGlyGly	274	
QY	1025	CACAGAAGTCTGGTGAATAAATGGACAAACATTCCTCAAAGCTCGTGTGATTTGCTCAGTG	1084	

DB	275	GlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerVal	294	
QY	1085	CCAGTCCAAATGCGCATTTGACACTCACTTTTATGAACTGCAGGATGTTCTTATGAAC	1144	
DB	295	ProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeuSer	313	
QY	1145	TTTAAAGATCCTAAAAATCCAGTTGTATATGAGGTGTTTACGACTTCAGATAACATTTTC	1204	
DB	314	SerArgAspArgGlnThrProLeuLeuTyrAlaValPheSerThrSerSerGlyValPhe	333	
QY	1205	AAGGATCAGCCGTGTGTATATAGCATGATGTGAGAGGGTGTTCCTTGGTCCA	1264	
DB	334	GlnGlySerAlaValCysValTyrSerMetAsnAspValArgAlaPheLeuGlyPro	353	
QY	1265	TATGCCACAGGATGAGCCCACTATCAATGGTGGCTTATCAAGGAGAGTCCCTTAT	1324	
DB	354	LeuProHisLysGluGlyProThrHisGlnTrpValSerTyrGlnGlyArgValProTyr	373	
QY	1325	CCAGGCCAGAACTTGTCCAGCAAAACATTTGTGTGTTTGTACTCTACAAAGACCTT	1384	
DB	374	ProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAspPhe	393	
QY	1385	CCTGATGATGTTATAACCTTTTCAAGAGTCCATCCAGCCATGTACAATCCAGTGTTCCT	1444	
DB	394	ProAspAspValIleGlnPheGlyArgAsnHisProLeuMetTyrAsnProValLeuPro	413	
QY	1445	ATGAACAATCCCAATAGTATCAAAACGGATGTAAATATCAATTTTACACAAATGTC	1504	
DB	414	MetGlyGlyArgProLeuPheLeuGlnValGlyAlaGlyTyrThrPheThrGlnIleAla	433	
QY	1505	GTAGACCGAGTGGATGACAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGT	1564	
DB	434	AlaAspArgValAlaAlaAspLysHisTyrAspValLeuPheIleGlyThrAspVal	453	
QY	1565	GGGACCGTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTAGAAGAGTT	1624	
DB	454	GlyThrValLeuLysValIleSerValProLysGlyArgArgProAsnSerGluGlyLeu	473	
QY	1625	CTGCTGGAAGAAATGACAGTGTTCGGGAACCGACTGCTATTTTACGCAATGGAGTTCC	1684	
DB	474	LeuLeuGluGluLeuGlnValPheGluAspSerAlaAlaIleThrSerMetGlnIleSer	493	
QY	1685	ACTAAGCAGCAACATATATGTTCAACGGTGGGTTGCCAGTCCCTCCCTTTTACAC	1744	
DB	494	SerLysArgGlnGlnLeuTyrValAlaSerArgAlaAlaValAlaGlnIleAlaLeuHis	513	
QY	1745	CGGTGTGATATTTACGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGT	1804	
DB	514	ArgCysThrAlaLeuGlyArgAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys	533	
QY	1805	GCTTGGGATGTTCTGCATGTTCTCGTATTTTCCACTGCAAGAGAGCCACAGCA	1864	
DB	534	AlaTrpAspGlySerAlaCysThrArgGlnProThrAlaLysArgArgPheArgArg	553	
QY	1865	CAAGATATAAGAAATGGAGACCCCTGACTCACTTCTCAGACTTACACCATGATATCAC	1924	
DB	554	GlnAspIleArgAsnGlyAspProSerThrLeuCysSerGly-----AspSerSer	570	
QY	1925	CATGCCACAGCCCTTAAGAGAGAAATCATCTATGTTGTGTAGAGAAATAGTACATTTTG	1984	
DB	571	HisSerValLeuLeuGluLysLysValLeu---GlyValGluSerGlySerAlaPheLeu	589	
QY	1985	GAATGCACTCCAGTCCAGAGCGCTGCTTATTTGGCAATTCAGAGCGGGAATGAA	2044	
DB	590	GluCysGluProArgSerLeuGlnAlaHisValGlnTrpThrPheGlnGlyAlaGlyGlu	609	
QY	2045	GAGCGAAAGAGAGATCAGATGATCATATCATCAGCAGACAGATCAAGCCCTTCG	2104	
DB	610	AlaAlaHisThrGlnValLeuAlaGluGluArgValGluArgThrAlaArgGlyLeuLeu	629	
QY	2105	CTACGTACTTACAAACAGAGGATTTACGCAATTTACCTCTGCGATCGGTGGAAATGGG	2164	

Db 445 ArgTyrGluValLeuPheLeuGlyThrAspArgGlyThrValGlnLysValLeu 464
 QY 1592 CTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAATGACAGCTTTTCGG 1651
 Db 465 ProLysAspAsp---GlnGluMetGluGluLeuMetGluGluValGluValPheLys 483
 QY 1652 GAACGACTGCTATTTCACCAATGGAGCTTCCATCAACGACCAACACTATATGGT 1711
 Db 484 AspProAlaProValLysThrMetThrIleSerSerLysArgGlnGlnLeuTyrValAla 503
 QY 1712 TCAACGGCTGGGTTCCCGAGCTCCCTTTACACCGGTGATATTTACGGGAAGCGTGT 1771
 Db 504 SerAlaValGlyValThrHisLeuSerLeuHisArgCysGlnAlaTyrGlyAlaCys 523
 QY 1772 GCTGAGTGTGCTCCCGAGACCCCTTACTGCTGCTGGATGGTTCCTGCTCGC 1831
 Db 524 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyGlnAlaCysSerArg 543
 QY 1832 TATTTTCCCACTGCAAGAGACGACACAGACGACAGATATAGAAATGGACCCACTG 1891
 Db 544 TyrThrAlaSerSerLysArgSerArgArgGlnAspValArgHisGlyAsnProIle 563
 QY 1892 ACTCACTGTTACAGACTTACACCATGATATACCATGGCCACAGCCCTGACAGAGAATC 1951
 Db 564 ArgGlnCysArgGlyPhe-----AsnSerAsnAlaAsnLysAsnAlaValGluSerVal 581
 QY 1952 ATCTATGTTGAGAAATAGTAGACATTTTGGAAATGAGTCCGAGTCCGAGAGAGG 2011
 Db 582 GlnTyrGlyValAlaGlySerAlaAlaPheLeuGluCysGlnProArgSerProGlnAla 601
 QY 2012 CTGCTATTGCAATTCAGAGGCGCAATGAAGACGAAAGACAGATCAGATGGAT 2071
 Db 602 ThrValTyrTrpLeuPheGlnArgAspProGlyAspArgArgGlnLeuAlaGlu 621
 QY 2072 GATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTACAACAGAGATTCA 2131
 Db 622 AspArgPheLeuArgThrGluGlnGlyLeuLeuLeuArgAlaLeuGlnLeuSerAspArg 641
 QY 2132 GGCATTTACTCTGCGATCGGTGGAAACATGGGTTTCATCAAACTCTTCTTAAGTAAAC 2191
 Db 642 GlyLeuTyrSerCysThrAlaThrGluAsnAsnPheLysHisValValThrArgValGln 661
 QY 2192 CTGGAATCATTCACACAGACATTTGGAGAACTTCTTCATAAGATCATGATGGAGAT 2251
 Db 662 LeuHisValLeuGlyArgAspAlaValHisAlaAlaLeuPhe----- 675
 QY 2252 GGTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAGGCTCTGG----- 2305
 Db 676 -----ProLeuSerMetSerAlaProProProGlyAlaGlyProPro 691
 QY 2306 -----TACAGACTTCATGACGCTCATCAACCCCAATCTCAACGATGGAT 2356
 Db 692 ThrProProTyrGlnGluLeuAlaGlnLeuLeuAlaGlnProGluValGlyLeuLeuHis 711
 QY 2357 GAGTCTGTGCAACACTTGGAAAGGACCGGAAACACAGCTCGGCAAGGCGACAT 2416
 Db 712 GlnTyrCysGlnGlyTyrTrpArgHisValProProSerProArgGluAlaProGly---- 730
 QY 2417 ACCCCAGGGAACAGTAACAAATGGAAGCACTTACAAGAAAAATAAGAAAGGTAGAACAGG 2476
 Db 731 -----AlaProArgSerProGluProGlnAspGlnLysLysProArgAsnArg 746
 QY 2477 AGGACCCAC 2485
 Db 747 ArgHisHis 749

RESULT 8

148748

semaphorin E - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I48748

R:Puschel, A.W.; Adams, R.H.; Betz, H.

Neuron 14, 941-948, 1995
 A>Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
 A:Reference number: I48744; MUID:95267431; PMID:7748561
 A:Accession: I48748
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-751 <RES>
 A:Cross-references: EMBL:X85994; NID:9854331; PIDN:CAA59986.1; PID:9854332
 C:Genetics:
 A:Gene: semE
 C:Superfamily: semaphorin

Alignment Scores: 3,31e-132 Length: 751
 Pred. No.: 1850.50 Matches: 343
 Score: 65.54% Conservative: 161
 Percent Similarity: 44.60% Mismatches: 224
 Best Local Similarity: 37.80% Indels: 41
 Query Match: 2 Gaps: 10
 DB:

US-09-774-490-1 (1-2709) x I48748 (1-751)

QY 215 AGGATTGCTCTCTTTCTGGGAGTATTACTTACAGCAAGCAAACTATCAGATGGG 274
 Db 4 ArgAlaIleCysValLeuValGlyValPheIleCysSerIleCys---ValArgGlySer 22
 QY 275 AAGAACATGTGCCAGCTGAATATTATCTACAAAGAAATGTTGGATCCAACTATGTG 334
 Db 23 SerGlnProGlnAlaArgValTyrLeuThrPheAspGluLeuArgGluThrLysThrSer 42
 QY 335 ATCACTTTCAATGGCTGGCCCAACAGCTCCAGTTATCATCTCTCTTTGGATCAGGAA 394
 Db 43 GluTyrPheSerLeuSerHisGlnGlnLeuAspTyrArgIleLeuLeuMetAspGluAsp 62
 QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTTCATTCGACCTGGTTAATATC 454
 Db 63 GlnAspArgIleTyrValGlySerLysAspHisIleLeuSerLeuAsnIleAsnIle 82
 QY 455 ---AAGGATTTTCAAAAGATTGTGGCCATGTTTACACAGAGAGATGAATGCAAG 511
 Db 83 SerGlnGluProLeuSerValPheTrpProAlaSerThrIleLysValGluGluCysLys 102
 QY 512 TGGGCTGGAAGAAGACATCTCTGAAGAATGCTGAATTTTCATCAAGGTACTTAAAGCATAT 571
 Db 103 MetAlaGlyLysAspProThrHisGlyCysGlyAsnPheValArgValIleGlnThrPhe 122
 QY 572 AATCAGACTCTCTGTAGCCTGTGGAAACGGGGCTTTTTCATCCAAATTTGCACCTACATT 631
 Db 123 AsnArgThrHisLeuTyrValCysGlySerGlyAlaPheSerProValCysThrTyrLeu 142
 QY 632 GAAATTGACATCATCTCTGAGACATATTTTAAAGCTGGAGAACTCACATTTTGAAGAAC 691
 Db 143 AsnArgGlyArgSerGluAspGlnValPheMetIleAsp---SerLysCysGluSer 161
 QY 692 GSCCTGGGAAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGA 751
 Db 162 GlyLysGlyArgCysSerPheAsnProAsnValAsnThrValSerValMetIleAsnGlu 181
 QY 752 GAAATTATCTCTGGAACCTGCACTGAATTTTATGGGCGAGACTTTTGTCTATCTTCCGAACT 811
 Db 182 GluLeuPheSerGlyMetTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgSer 201
 QY 812 CTGGGCAACCAACCCCAATCAGGACAGACGATGATTCAGGTGGCTCATATGATCCA 871
 Db 202 LeuThrLysArgMetGlnLeuArgThrAspGlnHisAsnSerLysTrpLeuSerGluPro 221
 QY 872 AAGTTCATTAGTGGCCACTCATCTCAGAGAGTGAACAATCTTGAAGATGACAAAGTATAC 931
 Db 222 MetPheValAspAlaHisValIleProAspGlyThrAspProAsnAspAlaLysValTyr 241
 QY 932 TTTTCTCTCCGTGAAATGCAATAGATGAGAACACTCTGTGAAAAGCTACTCAGCTAGA 991
 Db 242 PhePhePheLysGluArgLeuThrAspAsnAsnArgSerThrLysGlnIleHisSerMet 261

Qy	2069	GATGATCATATCATCAGGACAGATCAAGGCTCTTCTGCTACGTAGTCTATCAACAGAGAGGAT	2122
Db	618	AsnGluArgIleIleAlaThrSerGlnGlnGlyLeuLeuIleArgSerValGlnAspSerAsp	637
Qy	2129	TCAGGCAATTTACCTCTGCCATGCGGTGGACATGGGTTCATACAAACTCTCTTCTTAAGGTA	2188
Db	638	GlnGlyLeuThrHisCysIleAlaThrGluAsnSerPhelysGlnIleAlaLysIle	657
Qy	2189	ACCTTGGAAGTCAATTCACACAGCATTTTGGAAAGAACTTCTTCATAAAGATGATGGGA	2248
Db	658	AsnPhelysValLeuAspSer-----	664
Qy	2249	GATGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTAGCAGAGAGTCTCGTAC	2308
Db	665	-----GluMetValAlaValThrAspLysTrpSerProTrpThr	678
Qy	2309	-----AGAGACTTCATGTCAGCTCATCAAC	2332
Db	679	TrpAlaGlySerValArgAlaLeuProPheHisProLysAspIleLeuGlyAlaPheSer	698
Qy	2333	CACCCCAATCTCAACACGATGTGATGATCTGTGGAACAAAGTTTGGAAAGGACCGGAAAA	2392
Db	699	HisSerGluMetGlnLeuIleAsnGlnTyrCysLysAspThr-----ArgGlnGlnGln	716
Qy	2393	CAACGTGCGCAAGGCCAGGACATACCCAGGGAACAGTAAACAANTGGAAGCACTTACAA	2452
Db	717	GlnLeuGlyGlnGluProGlnLysMetArgGlyAspTyrGlyLysLeuLysAlaLeuIle	736
Qy	2453	GAATAATAAGAAAGGTAGAACAGGAGG	2479
Db	737	AsnSerArgLysSerArgAsnArgArg	745

RESULT 9

S66498

M-sema F protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999

C:Accession: S66498

R:Inagaki, S.; Furuyama, T.; Iwahashi, Y.

FEBS Lett. 370, 269-272, 1995

A:Title: Identification of a member of mouse semaphorin family.

A:Reference number: S66498; MUID:95385809; PMID:7656991

A:Accession: S66498

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-834 <NA>

C:Cross-references: EMBU:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599

C:Superfamily: semaphorin

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-834/Product: M-sema F protein #status predicted <MAT>

Alignment Scores:			
Pred. No.:	4,28e-65	Length:	834
Score:	966.00	Matches:	236
Percent Similarity:	51.35%	Conservative:	107
Best Local Similarity:	35.33%	Mismatches:	269
Query Match:	19.73%	Indels:	56
DB:	2	Gaps:	21

US-09-774-490-1 (1-2709) x S66498 (1-834)

Qy	206	TGCTTAAGTGGAGTCTGCTCTTTCTGGGAGTATTACTTACAGCAAGACAACTAT	265
Db	8	TrpLeu-----LeuAlaAlaGlyLeuTrpGlyLeuGlyIleGlyAlaGluMetTrpTrp	25
Qy	266	CAGAATGGGAAGCAAAATGTGCCAAGCTCAAAATTCCTTACAAAGAAATGTTTGGAAATCC	325
Db	26	-----AsnLeuValProArgLysThrValSerSerGlyGluLeuValThrVal	41
Qy	326	AACAATGTGATCACTTTCATGGCTTGGCCACAGCTCAGTATTCATCTTCCTTTTG	385
Db	42	ValArgPheSerGlnThrGlyIle-----GlnAspPheLeuThrLeu	58

Percent Similarity:	49.77%	Conservative:	114
Best Local Similarity:	32.63%	Mismatches:	261
Query Match:	17.51%	Indels:	73
DB:	2	Gaps:	23

US-09-774-490-1 (1-2709) x 148746 (1-782)

Qy	356	AACAGCTCCAGTTATATCATACCTCTCTTTTGGATGAGAACGGAGTAGGCTGTATGTTGGA	415
Db	12	AsnIleSerAenTyThrAlaLeuLeuLeuSerGlnAspGlyLysThrLeuTyrValGly	31
Qy	416	GCAAGGATCACATATTTTCATTTCGACCTGGTTATATCAG-----GAT	460
Db	32	AlaArgGluAlaLeuPheAlaLeuAsn---SerAenLeuSerPheLeuProGlyGlyGlu	50
Qy	461	TTTCAAAAGATTGTGTGCCAGTATCTTACACCACGAGAGATGAATGCAAGTGGCTCGGA	520
Db	51	TyrGlnGluLeuLeuTyrSerAlaAspAlaAspArgIysGlnGlnCysSerPheIysGly	70
Qy	521	AAAGACATCTCTGAAAGAATGTGCTAATTCATCAAGGTACTTAAGGCATATAATCAGACT	580
Db	71	LysAspProIysArgAspCysGlnAenTyIleIysIleLeuLeuProLeuAsnSerSer	90
Qy	581	CACTTGTACGCTCTGGNAACGGGGCTTTTCATCCAATTTCGACCTACATTGGAATTGGA	640
Db	91	HisLeuLeuThrCysGlyThrAlaAlaPheSerProLeuCysAlaTyIleHisIleAla	110
Qy	641	CATCATCTGTAGGACAATATTTTAAAGCTGGAG-----AATCATCATTTT	685
Db	111	Ser-----PheThrLeuAlaGlnAspGluAlaGlyAsnValIleLeu	124
Qy	686	GAATAACGGCGGTGGAAAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAATA	745
Db	125	GluAspGlyLysGlyHisCysProPheAspProAsnPheLysSerThrAlaLeuValVal	144
Qy	746	GATCGAATATTAATCTCTGNACTGCGAGCTGATTTTATGGGGCAGACTTTGCTATCTTC	805
Db	145	AspGlyGluLeuTyrThrGlyThrValSerSerPheGlnGlyAsnAspProAlaIleSer	164
Qy	806	CGAACTCTTGGGCACCAACCCACATGAGGACAGAGCATGATTCACAGTGGCTCAAT	865
Db	165	ArgSerGlnSer---SerArgProThrLysThrGlnSer---SerLeuAsnTrpLeuGln	182
Qy	866	GATCCAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGT---GACAATCTT-----GAA	916
Db	183	AspProAlaPheValAlaSerAlaThrSerProGluSerLeuGlySerProIleGlyAsp	202
Qy	917	GATGACAAAGTATACTTTTCTTCGTGGAAAATGCAATAGATGAGAACACTCTGGAAAA	976
Db	203	AspAspLysIleTyrPhePhePheSerGluThrGlyGlnGluPheGluPheGluAsn	222
Qy	977	GCTACTCACGCTAGAAATAGGTGCAGATGTCGAAGATGACTTTGAGGGCACAGAGTCTG	1036
Db	223	ThrIleValSerArgValAlaArgValCysLysGlyAspGluGlyGlyGluArgValLeu	242
Qy	1037	GTGAATAAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCACAAAT	1096
Db	243	GlnGlnArgTrpThrSerPheLeuLysAlaGlnLeuLeuCysSerArgPro---AspAsp	261
Qy	1097	GGCATTGACACTCATTTTGTGAACCTGCAGATGTATTCTCTAATG-----AATTTTAA	1150
Db	262	GlyPhePro-----PheAsnValLeuGlnAspValPheThrLeuAsnProAsnProGln	279
Qy	1151	GATCCTAAAAATCCAGTTGTATATGAGGTGTTTACGACTTCC-----AGTAACATTTC	1204
Db	280	AspTrpArgLysThrLeuSerIleGlyValPheThrSerGlnTrpHisArgGlyThrThr	299
Qy	1205	AAGGGATCAGCGGTGTGTATGATAGCATGAGTGTGAGAGGGGTGTTCTTCTTGGTCCA	1264
Db	300	GluGlySerAlaIleCysValPheThrMetAsnAspValGlnLysAlaPheAspGlyLeu	319
Qy	1265	TATGCCACAGGGATGGACCCAACTATCAATGGGTGCTTATCAAGGAAGATCCCTAT	1324

RESULT 11
I48745
semaphori

Db 558 GlnSerProProGlnLeuileLysGluValLeuThrValProAsnSerIleLeuGluLeu 577

QY 1211 TCAGCGGTGTGTATGATAGATGATGAGAGGGTGTCTCTGGTCCATATGCC 1270
Db SerAlaValCysAlaPheSerMetLysSerIleLeuGluSerPheAspGlyProPheLys 333
QY 1271 CACAGGATGGACCAACTATCAATGGTGGCTTATCAAGGA--AGAGTCCCTTATCCA 1327
Db GluGlnGluThrMetAsnSerAsnTrpLeuAlaValProSerLeuLysValProGluPro 353
QY 1328 CGGCAGGAACTGTGCCCAAGCAAAACATTTGGTGGTGTCTACACAAAGGACCTTCT 1387
Db ArgProGlyGlnCys-----ValAsnAspSerArgThrLeuPro 366
QY 1388 GATGATGTTATACCTTTGCAAGAGATGATCCAGCCATGTACAATCCAGTGTCTTANG 1447
Db AspValSerValAsnPheValLysSerHisThrLeuMetAspGluAlaValProAlaPhe 386
QY 1448 AACATCGCCCAATAGTCAAAACGGATGTAATATCAATTTACACAAATGTCCTA 1507
Db PheThrArgProIleLeuLeuLeuLeuLeuSerLeuGlnTrpArgPheThrLysIleAlaVal 406
QY 1508 GAC---CGAGTGGATGCAAGATGACAG---TATGATGTTATGTTTATCGAACACAT 1561
Db AspGlnGlnValArgThrProAspGlyLysAlaTyAspValLeuPheIleGlyThrAsp 426
QY 1562 GTTGGACCGTCTTAAAGTAGTT---TCAATTCCTAAGGAGACTTGGTATGATTAGAA 1618
Db AspGlyLysValIleLysAlaLeuAsnSerAlaSerPheAspSerSerAspThrValAsp 446
QY 1619 GAGGTTCCTGCGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTCA----- 1669
Db SerValIleGluGluLeuGlnValLeuProProGlyValProValLysAsnLeuTyr 466
QY 1670 GCAATGAGCTTCCACTAAGCAGCAACAACTATATATTTGGTTCAACGCTGGGGTTGCC 1729
Db ValValArgMetAspGlyAspAspSerLysLeuValValSerAspAspGluIleLeu 486
QY 1730 CAGCTCCCTTTACACCGGTGTGATATTTACGGGAAA-----GCGTGTGCTGAG 1777
Db AlaIleLysLeuHisArgCys-----GlySerAspLysIleThrAsnCysArgGlu 503
QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGCTGGATGCT----- 1816
Db CysValSerLeuGlnAspProTyrCysAlaTrpAspAsnValGlnLeuLysCysThrAla 523
QY 1817 -----TCTGATGTTCTCGCTATT----- 1837
Db ValGlySerProAspTrpSerAlaGlyLysArgPheIleGlnAsnIleSerLeuGly 543
QY 1837 ----- 1837
Db GluHisLysAlaCysGlyArgProGlnThrGluIleValAlaSerProValProThr 563
QY 1838 ---CCACTGCAAGAGACGCAAGACGACAGATATAGAAATGGAGACCCACTGACT 1894
Db GlnProThrThrLys-----SerSerGlyAspProVal--- 574
QY 1895 CACTGTTTCAGACTTACACCATGATAATCACCATGGCCAC---AGCCCTGAA----- 1945
Db HisSerIleHisGlnAlaGluPheGluProGluLeuAspAsn 588
QY 1946 AGAATCATCTATGTTGTTAGAGATAGT-----AGCACATTTTTCGAATGCACT 1993
Db GluIleValIleGlyValAspAspSerAsnValIleProAsnThrLeuAlaGluIleAsn 608
QY 1994 -----CCGACTCGCAGAGCGCTG 2014
Db HisAlaGlySerLysLeuProSerSerGlnGluLysLeu 621

RESULT 14

B49423

semaphorin I - fruit fly (*Drosophila melanogaster*) (fragment)C/Species: *Drosophila melanogaster*

C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999

C/Accession: B49423

R/Klodkin, A.I.; Matthes, D.J.; Goodman, C.S.

Cell 75, 1389-1399, 1993

A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone

A/Reference number: A49423; MUD:94094332; PMID:8269517

A/Accession: B49423

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-656 <KOL>

A/Cross-references: GB:L26082

C/Genetics:

A/Gene: semaI

A/Cross-references: FlyBase:FBgn0011259

Alignment Scores:			
Pred. No.:	2,53e-47	Length:	656
Score:	731.50	Matches:	199
Percent Similarity:	46.92%	Conservative:	129
Best Local Similarity:	28.47%	Mismatches:	248
Query Match:	14.94%	Indels:	123
DB:	2	Gaps:	24

US-09-774-490-1 (1-2709) x B49423 (1-656)

QY 536 GAATGTGCTAATTTTCATCAAGGTACTTAAGCATATATATCAGACTCACTTGTACGCTGT 595
Db 3 AspCysGlnAsnTyIleArgIleMetValValProSerProGlyArgLeuPheValCys 22
QY 596 GGAACGGGGGCTTTTCATCCAAATTTGC---ACCTACATTGAANTTGGACATCATCTGAG 652
Db GlyThrAsnSerPheArgProMetCysAsnThrTyIle-----IleSer 37
QY 653 GACAATATTTTAAGCTGGAGAACTCACATTTTGAACACGGCGTGGGAAGTCCATAT 712
Db 38 AspSerAsnTyThrLeuGluAlaThr-----LysAsnGlyGlnAlaValCysProTyr 55
QY 713 GACCTAAGCTGCTGACAGCATCCCTTTTAATAGATGGAGAATTATATCTCGGAACCTGCA 772
Db 56 AspProArgHisAsnSerThrSerValLeuAlaAspAsnGluLeuTyrSerGlyThrVal 75
QY 773 GCTCATTTTATGGGGCGAGACTTTGCTATCTTCGAACTCTTGGGCACCACCCCAATC 832
Db AlaAspPheSerGlySerAspProIleIleTyrArg-----GluProLeu 90
QY 833 AGGACAGAGCATGATTTCCAGGTGGCTCAATCATCAAAAGTTTCATTAGTGCCACCTC 892
Db 91 GlnThrGluGlnTyrAspSerLeuSerLeuAsnAlaProAsnPheValSerPhe--- 109
QY 893 ATCTCAGAGATGACAATCTCTGAAGATCACAAGATATACCTTTTCTCCGTGAAAATGCA 952
Db 110 -----ThrGlnGlyAspPheValTyrPhePheArgGluThrAla 123
QY 953 ATAGATGAGAGAACTCTGGAAGAGCTACTCACCTAGATAGTGCATGATGCAAGAT 1012
Db 124 ValGluPheIleAsnCysGlyLysAlaIleTyrSerArgValAlaArgValCysLysTrp 143
QY 1013 GACTTTGGGGGACACAGAGTCTGGTGAATAAATGGACAACATTCCTCAAAAGCTGCTG 1072
Db 144 AspLysGlyGlyProHisArgPheArgAsnArgTrpThrSerPheLeuLysSerArgLeu 163
QY 1073 ATTTGCTCAGTGGCAGGTCCAAATGGCATTCACACTCATTTTGTATGAACATGAGATGTA 1132
Db 164 AsnCysSerIleProGly-----AspTyrProPheTyrPheAsnGluIleGlnSerAla 181
QY 1133 TTCCTAATG-----AACTTTAAAGATCCTAAAATCCAGTTGTATATATGAGTGT 1183
Db 182 SerAsnLeuValGluGlyGlnTyrGlySerMetSerSerLysLeuIleTyrGlyValPhe 201
QY 1184 ACGACTTCCAGTACATTTTCAAGGGATCAGCGGTGTGTATGTATAGATGAGTGTG 1243
Db 202 AsnThrProSerAsnSerIleProGlySerAlaValCysAlaPheAlaLeuGlnAspIle 221
QY 1244 AGAAGGTGTTCCTTGTGTCATATGCCACAGGATGGACCCAACTACTCAATCAATGGTGCCT 1303

```
Db      222 AlaAspThrPheGluGlyGlnPheLysGluGlnThrGlyLeuAsnSerAsnTrpLeuPro 241
QY      1304 TAT---CAAGGAAGAGTCCCTTATCCAGCGCCAGGAATGTGTCGCCAGAAACATTGGT 1360
Db      242 ValAsnAsnAlaLysValProAspProArgProGlySerCys----- 255
QY      1361 GGTTCCTGCTTACAAAGACCTTCTGATGATGTTATTAACCTTTCCAGAGTCAATCCA 1420
Db      256 ---HisAsnAspSerArgAlaLeuProAspProThrLeuAsnPheIleLysThrHisSer 274
QY      1421 GCCATGTACAATCCAGTGTCTTATGAACAATCGCCCAATAGTATGATCAAAACCGATGTA 1480
Db      275 LeuMetAspGluAsnValProAlaPhePheSerGlnProIleLeuValArgThrSerThr 294
QY      1481 AATTATCAATTTACAAATTCCTGAGAC---CGAGTGGATGCAAGAGATGACAG--- 1534
Db      295 IleTyrArgPheThrGlnIleAlaValAspAlaGlnIleLysThrProGlyGlyLysThr 314
QY      1535 TATGATGTTATCTTATCGGAACAGATGTTGGACGGTCTTAAAGTAGTTCATTCCT 1594
Db      315 TyrAspValIlePheValGlyThrAspHisGlyLysIleIleLysSerValAsnAlaGlu 334
QY      1595 AAGGAGACTTGTATGAT---TTAGAAGAGGTCTGCTGGAAGAATGACAGCTTTTCGG 1651
Db      335 SerAlaAspSerAlaAspLysValThrSerValValIleGluGluLeuAspValLeuThr 354
QY      1652 GAACCGACTGTGCTTATTCAGCAATGAGCTTCCATTAAGACACAA----- 1696
Db      355 LysSerGluProIleArgAsnLeuGluIleValArgThrMetGlnTyrAspGlnProLys 374
QY      1697 -----CAACTATATATGTTCAACGGCTGGGGTCCCGAGCTC 1735
Db      375 AspGlySerTyrAspAspGlyLysLeuIleValThrAspSerGlnValValAlaIle 394
QY      1736 CCTTTACACCGGTGT---CATATTACGGGAAGCGTGTGCTGAGTGTGCTCCGCCGA 1792
Db      395 GlnLeuHisArgCysHisAsnAspLysIleThrSerCysSerGluCysValAlaLeuGln 414
QY      1793 GACCTTACTGTGCTGGATGGTCTGCA-----TGT----- 1825
Db      415 AspProTyrCysAlaTrpAspLysIleAlaGlyLysCysArgSerHisGlyAlaProArg 434
QY      1826 -----TCTCGTATTTT-----CCC 1840
Db      435 TrpLeuGluGluAsnTyrPheTyrGlnAsnValAlaThrGlyGlnHisAlaAlaCysPro 454
QY      1841 ACTGCAAGAGACGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1900
Db      455 SerGlyLysIleAsnSerLys-----AspAlaAsnAlaGlyGluGlyLysGlyPheArg 472
QY      1901 TCAGACTTACCATGATATACCATGCGCCACGCTGGAAGAGAGATCATCATGCT 1960
Db      473 AsnAspMetAspLeuLeuAspSerArgGlnSerLysAspGlnIleIleAspAsn 492
QY      1961 GTAGAGATAGTAGACATATTTTGGATCCAGTCCGAGTCCGAGAGCGCTGCTAT 2020
Db      493 IleAspLysAsnPheGlu-----GlyProGlnThrSerAlaAspIleIleAsn 508
QY      2021 TGGCAATTCAGAGCGCAATGAAGACGCAAGAAAGAGAGATCAGATGATCATCAT 2080
Db      509 AlaGlnTyr-----ThrValGluThrLeuVal 517
QY      2081 ATCAGACAGATCAGAGCCTTCTGCTACGATCTTACACAGAGAGATTCAGGCAATTAC 2140
Db      518 MetAlaValLeuAlaGlySerIlePheSerLeuLeuValGlyPhePheThrGlyThrPhe 537
QY      2141 CTC-----TGCCATCGGTGGAACATGGTTTCATACAACTCTTCTTAAGTAACC 2191
Db      538 CysGlyArgArgCysHisLysAspGluAspAsp-----Asn 549
QY      2192 CTGGAAGTCAATTCACACA-----GAGCATTTTGAAGAACTTCTTCATAAAGATGATGAT 2245
```

```
550 LeuProTyrProAspThrGluTyrPheGluGlnArgGlnAsnValAsnSerPhe 569
2246 GGAGATGGCTTAGACCAAAAGAAATGTCCTAATAGCATGACACCTAGCACAGAGTCTGG 2305
570 ProSerSerCysArgIleGlnGlnProLysLeuLeuProGlnValGluGluValThr 589
2306 TACAGAGACTTTCAGCTGAGCTCATCAACCCACCCCAATCTCAACACAGCTGATGATGCTGT 2365
590 TyrAlaGluProValLeuLeuProGlnProProProProAsnLysMet----- 605
2366 GAACAAGTTTGGAAAAAGGACCGAAACCAACGTCGGCAAGGCCAGGACATACCCAGGG 2425
606 -----HisSerProLys 609
2426 AACAGTACAAATCGAAGCACTTACAGAAATAAAGAGGTAGNAACAGAGGACC 2482
610 AsnThrLeuArgLysProProMetHisGlnMetHisGlnGlyProAsnSerGluThr 628

RESULT 15
JC5928
semaphorin F precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C:Accession: JC5928
R:Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candi
A:Reference number: JC5928; MUID:98125554; PMID:9464278
A:Accession: JC5928
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1074 <STM>
A:Cross-references: GB:U52840; NID:g2722583; PIDN:AAC09473.1; PID:g2722584
A:Experimental source: brain
C:Comment: This protein disrupts normal brain development and leads to some of the featu
C:Genetics:
A:Gene: semaf
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:50-533/Domain: semaphorin #status predicted <SEM>
F:840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F:971-993/Domain: transmembrane #status predicted <TM>

Alignment Scores:
Pred. No.: 3,598-47 Length: 1074
Score: 730.00 Matches: 185
Percent Similarity: 51.25% Conservative: 102
Best Local Similarity: 33.04% Mismatches: 205
Query Match: 14.91% Indels: 68
Gaps: 22

US-09-774-490-1 (1-2709) x JC5928 (1-1074)
QY 176 AAAGGACCTACAGCGTCTGCAGCATGGCTGGTGTAACTAGGATGTCTGCTTTCTGG 235
Db 2 LysGlyThr-----CysVallealatrPheSerSerLeuGlyLeuTrpArg 18
QY 236 GGAGTATTACTTACAGCAAGACGAAACTATCATGATGGGAAGAACATGTGCCAAGGCTG 295
Db 19 LeuAlaHisProGluAlaGlnGlyThrThrGlnCysGlnArgThrGluHisPro----- 36
QY 296 AAATTATCCTACAAGAAATG-----TTGGAATCCAACAATGTGATC 337
Db 37 VallieserTyrLysGluIleGlyProTrpLeuArgGluPheArgAlaLysAsnAlaala 56
QY 338 ACTTTCATGCTTGGCCACACAGCTCCAGTTATCATACCTTCCTTTTGGATGAGGAACGG 397
Db 57 AspPheSerGlnLeu-----ThrPhe-----AspProGlyGln 67
QY 398 AGTAGGCTGTATGTTGGAGCAAGATGACATATTTTCGACCTGGTTAATATCAAG 457
Db 68 LysGluLeuValGlyAlaArgAsnTyrLeuPheArgLeuGlnLeuGluAspLeuSer 87
```

458 QY GATTTTCAAAGATGTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGGCGT 517
 88 Db LeuileGlnalaValGluuTpgluCysAspGlualaThrLysLysAlaCysTyrSerLys 107
 518 QY GGAAGAAGCATCTCTGAAGAAGATGCTTAATTTTCATCAAGGTACTTAAAGGCATATATACAG 577
 108 Db GlyLysSer---LysGluGluCysGlnAsnTyrileArgValLeu---LeuValGlyGly 125
 578 QY ACTCACATTGACGCTCTGGAAACGGGGCTTTTTCATCAATTTTCACCTACATTTGAAAT 637
 126 Db AspArgLeuPheThrCysGlyThrAsnAlaPheThrProValCysThrAsnArgSerLeu 145
 638 QY GGACATCATCTGTAG-----GACATATATTTTAACTGGAGAACTCACATTTTGAAGAAC 691
 146 Db SerAsnLeuAlaGluHisAspGlnle-----Ser 156
 692 QY GGCCGTGGGAAGAGTCCATATGACCCCTAAG---CTGCTGACAGCATCCCTTTTAATAGAT 748
 157 Db GlyMetAlaArgCysProTyrSerProGlnHisAsnSerThrAlaLeuLeuThrAlaGly 176
 749 QY GGGAATATATACTCTGCAACTGTCAGCTGATTTTATGGGGCAGACTTTGCTATCTTCCGA 808
 177 Db GlyGluLeuTyrAlaAlaThrAlaMetAspPheProGlyArgAspProAlaIleTyrArg 196
 809 QY ACTCTTGGGCACCAACCCCAATCAGACAGACAGCATGATTCAGGTGGCTCAATGAT 868
 197 Db SerLeuGlyIleLeuProProLeuArgThrAlaGlnTyrAsnSerLysTyrTpLeuAsnGlu 216
 869 QY CCAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGCACATCCTGAAGATGACAAGTA 928
 217 Db ProAsnPheValSerSerTyrAspIleGly-----AsnPheThr 229
 929 QY TACTTTTCTTCCGTGAAATAGCATAGATGGAGAACACTCT---GGAAAAGCTACTCAC 985
 230 Db TyrPhePhePheArgGluAsnAlaVal-----GluHisAspCysGlyLysThrValPhe 247
 986 QY GCTAGAAATAGTCTAGATATCAAGAATGACTTTGGAGGGCAGACAAAGTCTCGTGAATAAA 1045
 248 Db SerArgAlaAlaArgValCysLysAsnAspIleGlyArgPheLeuLeuGluAspThr 267
 1046 QY TGGACACANTCTCAAGCTCGTCTGATTTGCTCACTGCCAGTCCAAATGGCATGAC 1105
 268 Db TrpThrThrPheMetLysAlaArgLeuAsnCysSerArgProGlyGlu-----ValPro 285
 1106 QY ACTCATTTTGCATGAACCTGGCAGGATGTATCTCATGAACCTTAAAGATCTCAAAAATCCA 1165
 286 Db PheTyrTyrAsnGluLeuGlnSerThrPhePheLeu-----ProGluLeuAsp 301
 1166 QY GTTGATATAGAGTGTTTAGCACTTCACAGTAACATTTTCAAGGATCAGCGCTGTGTATG 1225
 302 Db LeuileTyrGlyIlePheThrThrAsnValAsnSerileAlaAlaSerAlaValCysVal 321
 1226 QY TATAGCATGATGATGTGAGAAGGGTGTCTTGGTCCATATGCCCAAGGGATGGACCC 1285
 322 Db PheAsnLeuSerAlaIleAlaGlnAlaPheSerGlyProPheLysTyrGlnGluAsnSer 341
 1286 QY AACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTCATCCAGCGCCAGCACTTGTC 1345
 342 Db ArgSerAlaTrpLeuProTyrProAsnProAsnProHisPheGlnCysGlyThrValAsp 361
 1346 QY AGCAAAACATTTGGTGGTGTGTGACTCTCAAAAGGACCTCTCCTGATGATGTTATAACCTTT 1405
 362 Db GlnGlyLeuTyrValAsnLeu---ThrGluArgAsnLeuGlnAsp----- 375
 1406 QY GCAGAAGTATCCAGGCATGTACATCCAGTGTTCCTATGAACAATCCGCCAATAGTG 1465
 376 Db AlaGlnLysPheIleLeuValHisGluValValGlnProValThrThrValProSerPhe 395
 1466 QY ATCAAAACGGATGTAAATATTCAATTTACAAATTTCTCTAGACCGAGTGGATGCAGAA 1525
 396 Db MetGluAsp-----AsnSerArgPheSerHisValAlaValAspValValGlnGlyArg 413
 1526 QY GATGGAAGATGATGATGTATGTTTATCGGAACAGATGTTGGGACCGTTCCTTAAGTAGTT 1585

	:::	::::::::::::::					
Db	414	GluAlaLeuValHisIleIleTyrLeuAlaThrAspTyrGlyThrIleLysValArg	433				
Qy	1586	TCAATTCTAAGACAGACTTGATGATTAGAAGAGGTTCGTGGGAAGAAATCACAAGTT	1645				
		::::					:::
Db	434	ValProLeuAsnGlnThr-----Ser-Ser-SerCysLeuLeuGluGluIleGluLeu	450				
Qy	1646	TTTT-----CGGGAACCGACTGCTATTTCAGCAATGAGAGTTCCTCCACTAAGCAC	1693				
Db	451	PheProGluargargargGluPro-----IlEargSerLeuGlnIleLeuHisSerGln	468				
Qy	1694	CAACAATAATATTGGTTCACGGCTGGGGTGCACCAGCTCCCTTTACACCGGTGTGAT	1753				
Db	469	SerValLeuPheValGlyLeuArgGluHisValValLysIleProLeuLysArgCysGln	498				
Qy	1754	ATTTACGGGAARAGCGTGTGCTGTGAGTGTGCTCGCCGAGACCTTACTGTCTGGGAT	1813				
Db	489	PheTyr---ArgThrArgSerThrCysIleGlcValaGlnAspProTyrCysGlyVtPasp	507				

Search completed: August 3, 2003, 09:56:58
Job time : 90 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 31, 2003, 11:37:35 ; Search time 9579 Seconds
(without alignments)
11569.494 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatctttatttcatgatg.....agggttttttcttaataacc 2709

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2709	100.0	2709	6	AX207154	AX207154 Sequence
2	2508.4	92.6	2530	6	AX743092	AX743092 Sequence
3	2508.4	92.6	2530	6	BD084822	BD084822 Diagnosis
4	2508.4	92.6	2530	9	HUMHSEM	L26081 Homo sapien
5	2508.4	92.6	2530	11	G31703	G31703 SWSS1973 Er
6	2508.4	92.6	2601	6	AR040598	AR040598 Sequence
7	2508.4	92.6	2601	6	I47054	I47054 Sequence 53
8	2050	75.7	5952	10	MUSC1	BD5028 Mus musculus
9	2027.6	74.8	2913	10	MNRNASEMD	X85993 M.musculus
10	1909.4	70.5	2319	10	RNSIICNI	X95286 R.norvegicu
11	1883.8	69.5	2319	10	MUSCOSE	L41541 Mus musculu
12	1744	64.4	3263	5	GGU02528	U02528 Gallus gall
13	1612.4	59.5	1998	10	MUSSEMAIII	L40484 Mus musculu
14	1477	54.5	2325	5	AY030051	AY030051 Xenopus 1
15	1415.2	52.2	1481	6	I11900	I11900 Sequence 1
16	1069.2	39.5	2337	5	AF083382	AF083382 Danio rer
17	1035.8	38.2	3148	5	AF086761	AF086761 Danio rer
18	718.6	26.5	2331	6	E27342	E27342 Novel poly
19	718.6	26.5	3871	6	AX376280	AX376280 Sequence
20	718.6	26.5	3871	6	AX697241	AX697241 Sequence
21	718.6	26.5	3880	6	E27343	E27343 Novel poly
22	697.2	25.7	2715	5	GGU28240	U28240 Gallus gall
23	652	24.1	169408	2	AC073110	AC073110 Homo sapi
24	650.4	24.0	179640	9	AC006322	AC006322 Homo sapi
25	647.4	23.9	6474	6	AX743098	AX743098 Sequence
26	647.4	23.9	6474	9	AB002329	AB002329 Human mRN
27	615.6	22.7	2615	5	AF022947	AF022947 Gallus ga
28	599.4	22.1	2337	6	AX155172	AX155172 Sequence
29	596.4	22.0	2898	10	AF034744	AF034744 Mus muscu
30	595.8	22.0	2872	10	MNRNASEMA	X85990 M.musculus
31	594.8	22.0	2328	10	MNRSEMH	Z80941 M.musculus
32	593.8	21.9	2879	5	AF124485	AF124485 Danio rer
33	581.6	21.5	2919	9	AX743094	AX743094 Sequence
34	581.6	21.5	2919	9	HSU28369	U28369 Homo sapien
35	569.6	21.0	2813	9	AF217991	AF217991 Homo sapi
36	568.8	21.0	2809	9	BC024220	BC024220 Homo sapi
37	568.8	21.0	2825	9	BC013975	BC013975 Homo sapi
38	566.6	20.5	2300	9	AB083186	AB083186 Homo sapi
39	554.4	20.5	2975	6	BD106505	BD106505 Human sem
40	554.4	20.5	4996	9	BC030690	BC030690 Homo sapi
41	553.6	20.4	3988	10	NM293948	Z93948 M.musculus
42	553.6	20.4	4466	10	NM293947	Z93947 M.musculus
43	552.8	20.4	5177	6	AX743096	AX743096 Sequence
44	552.8	20.4	5177	9	AB000220	AB000220 Homo sapi
45	551.2	20.3	2336	5	AF022946	AF022946 Gallus ga

ALIGNMENTS

RESULT 1
AX207154
LOCUS AX207154 2709 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0155455.
ACCESSION AX207154
VERSION AX207154.1 GI:15394943
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS Jin, S.
TITLE Resistance sequences and uses thereof
JOURNAL Patent: WO 0155455-A 1 02-AUG-2001;

FEATURES		Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)	
source	Location/Qualifiers		
	1..2709		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
BASE COUNT	831 a 560 c 620 g 698 t		
ORIGIN			
	Query Match	100.0%;	Score 2709; DB 6; Length 2709;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 2709;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	AATCTTTTATTTATCGATGTTAAACAAGCTTACGATGCCACGTCGAGGGGTGTCG	60
Db	1	AATCTTTTATTTATCGATGTTAAACAAGCTTACGATGCCACGTCGAGGGGTGTCG	60
Qy	61	ACCACGCGTCGGGAGTAGTTGAGCTCGCCTGTTCTCCCATGTCAGCCAGTCTATTT	120
Db	61	ACCACGCGTCGGGAGTAGTTGAGCTCGCCTGTTCTCCCATGTCAGCCAGTCTATTT	120
Qy	121	CCAGATTGTTGAACTTCTCTGCGCCGCAATAACAGGAAGGAGACTAAAGCAGCAAGG	180
Db	121	CCAGATTGTTGAACTTCTCTGCGCCGCAATAACAGGAAGGAGACTAAAGCAGCAAGG	180
Qy	181	GACCTACAGCGTCTGCAGATGGGCTGTTAACTAGGATTGTCGTCTTTCTGGGGAGT	240
Db	181	GACCTACAGCGTCTGCAGATGGGCTGTTAACTAGGATTGTCGTCTTTCTGGGGAGT	240
Qy	241	ATTACTTACAGCAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
Db	241	ATTACTTACAGCAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
Qy	301	ATCCTACAAAGAAATGTTGGAATCCAACTAATGTGATCACTTTCAATGGCTTGGCCAAAG	360
Db	301	ATCCTACAAAGAAATGTTGGAATCCAACTAATGTGATCACTTTCAATGGCTTGGCCAAAG	360
Qy	361	CTCCAGTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAA	420
Db	361	CTCCAGTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAA	420
Qy	421	GGATCACATATTTTTCATTCGACCTGTTTAAATCAAGGAATTTCAAAAGATTTGTGGCC	480
Db	421	GGATCACATATTTTTCATTCGACCTGTTTAAATCAAGGAATTTTCAAAAGATTTGTGGCC	480
Qy	481	AGTATCTTACACGAGAGATGAATGCAAGTGGGCTGGAAGACATCTCTGAAAGAATG	540
Db	481	AGTATCTTACACGAGAGATGAATGCAAGTGGGCTGGAAGACATCTCTGAAAGAATG	540
Qy	541	TGCTAATTTTCATCAAGGTACTTAAAGGCATATATCAGACTCACTTGTAGCCCTGTGAAC	600
Db	541	TGCTAATTTTCATCAAGGTACTTAAAGGCATATATCAGACTCACTTGTAGCCCTGTGAAC	600
Qy	601	GGGGGCTTTTCATCCAAATTTGCACTTACATTTGAAATTTGACATCATCTCTGAGCAATAT	660
Db	601	GGGGGCTTTTCATCCAAATTTGCACTTACATTTGAAATTTGACATCATCTCTGAGCAATAT	660
Qy	661	TTTTAAGCTGGAGAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAA	720
Db	661	TTTTAAGCTGGAGAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAA	720
Qy	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATTTACTCTGGAACCTGCAGCTGATTT	780
Db	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATTTACTCTGGAACCTGCAGCTGATTT	780
Qy	781	TATGGGGGAGACTTTTGCTATCTTCCGAACCTCTTGGGCAACCAACCCCAATCAGGACAGA	840
Db	781	TATGGGGGAGACTTTTGCTATCTTCCGAACCTCTTGGGCAACCAACCCCAATCAGGACAGA	840
Qy	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGA	900
Db	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGA	900

QY	901	GAGTGACAAATCTCTGAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGG	960
DB	901	GAGTGACAAATCTCTGAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGG	960
QY	961	AGAACACTCTTGAAAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACTTTGG	1020
DB	961	AGAACACTCTTGAAAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACTTTGG	1020
QY	1021	AGGGCACAGAGTCTGCTGAAATGACACATCTCTCAAGCTCGTCTGATTTGCTC	1080
DB	1021	AGGGCACAGAGTCTGCTGAAATGACACATCTCTCAAGCTCGTCTGATTTGCTC	1080
QY	1081	AGTGCCAGCTCAAAATGCGCATTTGACACTCAITTTGATGAATGCAAGGATGATTTCTTAAT	1140
DB	1081	AGTGCCAGCTCAAAATGCGCATTTGACACTCAITTTGATGAATGCAAGGATGATTTCTTAAT	1140
QY	1141	GAACTTTTAAAGATCCTAAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCCAGTAACAT	1200
DB	1141	GAACTTTTAAAGATCCTAAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCCAGTAACAT	1200
QY	1201	TTTCAAGGGATCAGCCGTGTGTATGATAGCATGATGTGAGAGGGTGTTCCTTGG	1260
DB	1201	TTTCAAGGGATCAGCCGTGTGTATGATAGCATGATGTGAGAGGGTGTTCCTTGG	1260
QY	1261	TCCATATGCCACAGGGATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCC	1320
DB	1261	TCCATATGCCACAGGGATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCC	1320
QY	1321	CTATCCAGGCGCAGGAACCTTGTCCAGCAAAACAATTTGGTGGTGTGACTCTTACAAGA	1380
DB	1321	CTATCCAGGCGCAGGAACCTTGTCCAGCAAAACAATTTGGTGGTGTGACTCTTACAAGA	1380
QY	1381	CTTCTCTCATGATGTTATAA CTTTGGCAAGAGTCCAGCCATGATCAATCCAGTGT	1440
DB	1381	CTTCTCTCATGATGTTATAA CTTTGGCAAGAGTCCAGCCATGATCAATCCAGTGT	1440
QY	1441	TCCATGAAACAATCGCCCAATAGTATCAAAACGGATGTAATTAATCAATTTACACAAT	1500
DB	1441	TCCATGAAACAATCGCCCAATAGTATCAAAACGGATGTAATTAATCAATTTACACAAT	1500
QY	1501	TGTCGTAGACCGAGTGGATGCAAGAGATGGAACAGTATGATTTATGTTTATCGGAACAGA	1560
DB	1501	TGTCGTAGACCGAGTGGATGCAAGAGATGGAACAGTATGATTTATGTTTATCGGAACAGA	1560
QY	1561	TGTTGGGACCTCTTAAAGTAGTTTCAATTCCTAAGAGAGCTTGGTATGATTTAGAAGA	1620
DB	1561	TGTTGGGACCTCTTAAAGTAGTTTCAATTCCTAAGAGAGCTTGGTATGATTTAGAAGA	1620
QY	1621	GGTTCGTGGAAGAAATGACAGTGTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCT	1680
DB	1621	GGTTCGTGGAAGAAATGACAGTGTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCT	1680
QY	1681	TTCCAATAAGCAGCAACAACTATATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTT	1740
DB	1681	TTCCAATAAGCAGCAACAACTATATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTT	1740
QY	1741	ACACCGGTGATATTTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGAGAGCCCTTA	1800
DB	1741	ACACCGGTGATATTTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGAGAGCCCTTA	1800
QY	1801	CTGTGCTGGGATGGTCTGCTGATGTTCTGCTATTTTCCCACTGCAAAAGAGACGCAACAG	1860
DB	1801	CTGTGCTGGGATGGTCTGCTGATGTTCTGCTATTTTCCCACTGCAAAAGAGACGCAACAG	1860
QY	1861	ACGACAAGATATAAATAATGAGACCCACTGACTCACTGTTTCAAGACTTACACCATGATAA	1920
DB	1861	ACGACAAGATATAAATAATGAGACCCACTGACTCACTGTTTCAAGACTTACACCATGATAA	1920
QY	1921	TCACCATGGCCACAGCCCTGAGAGAGAAATCATCTATGCTGTAGAGATAGTAGCACATT	1980
DB	1921	TCACCATGGCCACAGCCCTGAGAGAGAAATCATCTATGCTGTAGAGATAGTAGCACATT	1980
QY	1981	TTTGGAAATGCAGTCCGAAAGTCGAGAGAGCGCTGCTCTATTGGCAATTTCCAGAGCGAAA	2040

913 TGAAGATGACAAAGTATACCTTTTCTTCCGTGAAATGCAATAGATGAGAAACACTCTCG 972
Db |||||
729 TGAAGATGACAAAGTATACCTTTTCTTCCGTGAAATGCAATAGATGAGAAACACTCTCG 788
Qy |||||
973 AAAAGCTACTCAGCTAGAAATAGGTGAGATATCAAGAATGACTTTGGAGGGACAGAAG 1032
Db |||||
789 AAAAGCTACTCAGCTAGAAATAGGTGAGATATCAAGAATGACTTTGGAGGGACAGAAG 848
Qy |||||
1033 TCTGGTGAATTAATGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 1092
Db |||||
849 TCTGGTGAATTAATGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 908
Qy |||||
1093 AAATGGCAATGACACTCAATTTGATGAACTGACAGATGATTTCTTAATGAACTTTAAAGA 1152
Db |||||
909 AAATGGCAATGACACTCAATTTGATGAACTGACAGATGATTTCTTAATGAACTTTAAAGA 968
Qy |||||
1153 TCCTAAATTCAGTTGTATATGAGTGTGTTACGACTTCCAGTAACATTTTCAAGGATC 1212
Db |||||
969 TCCTAAATTCAGTTGTATATGAGTGTGTTACGACTTCCAGTAACATTTTCAAGGATC 1028
Qy |||||
1213 AGCCGTGTATGTATAGATGATGATGAGAGGGTGTCTTGGTCCATATGCCA 1272
Db |||||
1029 AGCCGTGTATGTATAGATGATGATGAGAGGGTGTCTTGGTCCATATGCCA 1088
Qy |||||
1273 CAGGGATGGAACCCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCTATCCAGGCC 1332
Db |||||
1089 CAGGGATGGAACCCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCTATCCAGGCC 1148
Qy |||||
1333 AGAACTTGTCCAGCAAAACATTTGGTGTGTTGACTCTCAAAAGGACCTTCTGATGA 1392
Db |||||
1149 AGAACTTGTCCAGCAAAACATTTGGTGTGTTGACTCTCAAAAGGACCTTCTGATGA 1208
Qy |||||
1393 TGTATAA CTTTGGCAAGAGTATCCAGCATGTATCAATTCAGTGTCTTCTATGACAA 1452
Db |||||
1209 TGTATAA CTTTGGCAAGAGTATCCAGCATGTATCAATTCAGTGTCTTCTATGACAA 1268
Qy |||||
1453 TCGCCCAATAGTGATCAAAACCGATGTAAATTTATCAATTTACAAATTTGTCTGACCG 1512
Db |||||
1269 TCGCCCAATAGTGATCAAAACCGATGTAAATTTATCAATTTACAAATTTGTCTGACCG 1328
Qy |||||
1513 AGTGAATGCAAGATGGA CAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db |||||
1329 AGTGAATGCAAGATGGA CAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1388
Qy |||||
1573 TCTTAAGTAGTTTCAATTCCTAAGCAGACTGGTATGATTTAGAGAGTCTCTCTGGA 1632
Db |||||
1389 TCTTAAGTAGTTTCAATTCCTAAGCAGACTGGTATGATTTAGAGAGGTTCTCTGGA 1448
Qy |||||
1633 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA 1692
Db |||||
1449 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA 1508
Qy |||||
1693 GCAACAACTATATATTTGTTTCAACCGCTGGGTGCCCAGTCCCTTTTACACCGGTGGA 1752
Db |||||
1509 GCAACAACTATATATTTGTTTCAACCGCTGGGTGCCCAGTCCCTTTTACACCGGTGGA 1568
Qy |||||
1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGGA 1812
Db |||||
1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGGA 1628
Qy |||||
1813 TGCTTCTGATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1872
Db |||||
1629 TGCTTCTGATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1688
Qy |||||
1873 AGAAATGAGACCCCACTGACTGCTGTTTACAGCTTTACCAATGATAATCAATGGCCA 1932
Db |||||
1689 AGAAATGAGACCCCACTGACTGCTGTTTACAGCTTTACCAATGATAATCAATGGCCA 1748
Qy |||||
1933 CAGCCCTGAAGAGAGATCATCTATGCTGAGAGATGATGACATTTTTCGAATGAG 1992
Db |||||
1749 CAGCCCTGAAGAGAGATCATCTATGCTGAGAGATGATGACATTTTTCGAATGAG 1808

Qy 1993 TCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 2052
Db |||||
1809 TCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAGAGCGCTTCTGCTAGCTAG 2112
Db |||||
1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAGAGCGCTTCTGCTAGCTAG 1928
Qy 2113 TCTCAACACAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAAACATGGGTTCATACA 2172
Db |||||
1929 TCTCAACACAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAAACATGGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGSTAACTCCTGGAAGTCAITGACA CAGAGCATTTTGGAGAACTTCTTCA 2232
Db |||||
1989 AACTCTTCTTAAGSTAACTCCTGGAAGTCAITGACA CAGAGCATTTTGGAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2292
Db |||||
2049 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2108
Qy 2293 CCAGAAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACACCCCAATCTCAACACGAT 2352
Db |||||
2109 CCAGAAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTCTGCAACAAAGTTTGGAAAGGACCGAATAACAACTCGGCAAGGCCAGG 2412
Db |||||
2169 GGATGAGTCTCTGCAACAAAGTTTGGAAAGGACCGAATAACAACTCGGCAAGGCCAGG 2228
Qy 2413 ACATACCCCGGGAACAGTAACTTGAAGAGCATTACAGAAATAAGAAAGGTAGAAA 2472
Db |||||
2229 ACATACCCCGGGAACAGTAACTTGAAGAGCATTACAGAAATAAGAAAGGTAGAAA 2288
Qy 2473 CAGAGGACCCCAAGATTTGAGAGGACCCAGAGAGTCTGAGCTGCATTACCTCTAGA 2532
Db |||||
2289 CAGAGGACCCCAAGATTTGAGAGGACCCAGAGAGTCTGAGCTGCATTACCTCTAGA 2348
Qy 2533 AACCTCAAACTAGAGAACTTGGCTAGACAAATAA CTGGAATAAATAAATAATATACAT 2592
Db |||||
2349 AACCTCAAACTAGAGAACTTGGCTAGACAAATAA CTGGAATAAATAAATAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCATTTATGAGATGTTTACAAATGTTGGGAATTCAGCTGAGTTCCA 2652
Db |||||
2409 GAACTTTTTCATGGCATTTATGAGATGTTTACAAATGTTGGGAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTC 2702
Db |||||
2469 CCAATTATAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTC 2518

RESULT 3

BD084822 2530 bp DNA linear PAT 27-AUG-2002
Locus BD084822
DEFINITION Diagnosis method and reagents.
ACCESSION BD084822
VERSION BD084822.1 GI:22630432
KEYWORDS JP 2001522241-A/15.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Leeuwen, F.W.V., Grosveld, F.G. and Burbach, J.P.H.
Patent: JP 2001522241-A 15 13-NOV-2001;
Diagnosis method and reagents
TITLE ROTTERDAM ACADEMY OF ARTS AND SCIENCES, ERASMUS UNIVERSITY
JOURNAL ROTTERDAM ACADEMY OF UTRECHT
COMMENT OS Homo sapiens (human)
PN JP 2001522241-A/15
PD 13-NOV-2001
PF 02-APR-1998 JP 1998542545
PR 10-APR-1997 US 60/043163
PI FREDERIK W VAN LEEUWEN, FRANKLIN G GROSVELD, JOHANNES PETER PI
HENRI BURBACH

```
PC C12Q1/68,C07K14/47,C12N15/52,C12N9/00,C12N5/10,A61K38/43, PC
A01K67/027,
PC A01K48/00//C07K16/18
CC Strandedness: Double;
CC Topology: Linear;
CC senaphorin-III gene, Genbank accession number L26081 FH Key
CC Location/Qualifiers
FT source 1..2530
FT Location/Qualifiers
   /organism="Homo sapiens (human)".
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN
Query Match 92.6%; Score 2508.4; DB 6; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 193 CTGACGATGGGCTGTTAACTAGGATGCTGCTCTTTCTGGGAGTATTACTTACAGC 252
DB |||||
9 CTGACGATGGGCTGTTAACTAGGATGCTGCTCTTTCTGGGAGTATTACTTACAGC 68
DB |||||
QY 253 AAGAGCAAACTATCAGAATGGGAAGAACAAATGTCGAAGGCTGAAATTAATCTACAAAGA 312
DB |||||
69 AAGAGCAAACTATCAGAATGGGAAGAACAAATGTCGAAGGCTGAAATTAATCTACAAAGA 128
DB |||||
QY 313 AATGTTGGAAATCCAACTATGATCACTTTCAATGGCTGGGCAACAGCTCCAGTTATCA 372
DB |||||
129 AATGTTGGAAATCCAACTATGATCACTTTCAATGGCTGGGCAACAGCTCCAGTTATCA 188
DB |||||
QY 373 TACCTTCTCTTTGGATGAGGAACGAGTAGGCTGATGTTGGAGCAAGGATCACATATT 432
DB |||||
189 TACCTTCTCTTTGGATGAGGAACGAGTAGGCTGATGTTGGAGCAAGGATCACATATT 248
DB |||||
QY 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 492
DB |||||
249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 308
DB |||||
QY 493 CAGAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTCAT 552
DB |||||
309 CAGAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTCAT 368
DB |||||
QY 553 CAAGGTACTTAAAGCATATAATCAGACTCAGCTTGTAGCGCTGTGGAACGGGGCTTTTCA 612
DB |||||
369 CAAGGTACTTAAAGCATATAATCAGACTCAGCTTGTAGCGCTGTGGAACGGGGCTTTTCA 428
DB |||||
QY 613 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGA 672
DB |||||
429 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGA 488
DB |||||
QY 673 GAACTCAGATTTTGAACCGGCTGGGAAGAGTCATATGACCTTAAGCTGCTGACAGC 732
DB |||||
489 GAACTCAGATTTTGAACCGGCTGGGAAGAGTCATATGACCTTAAGCTGCTGACAGC 548
DB |||||
QY 733 ATCCCTTTTAAATAGATGGGAATTAATCTCTGGAATCTGAGCTGATTTTATGGGCGAGA 792
DB |||||
549 ATCCCTTTTAAATAGATGGGAATTAATCTCTGGAATCTGAGCTGATTTTATGGGCGAGA 608
DB |||||
QY 793 CTTTGTCTATCTTCCGAACTTTGGGCAACCAACCAATCAGGACAGAGCAGCATGATTC 852
DB |||||
609 CTTTGTCTATCTTCCGAACTTTGGGCAACCAACCAATCAGGACAGAGCAGCATGATTC 668
DB |||||
QY 853 CAGGTGGCTCAATGATCCAAAGTTCAATAGTGCCCACTCAATCTCAGAGAGTGAACAATCC 912
DB |||||
669 CAGGTGGCTCAATGATCCAAAGTTCAATAGTGCCCACTCAATCTCAGAGAGTGAACAATCC 728
DB |||||
QY 913 TGAAGATGACAAGTATATCTTTCTCGGTGAAAATGCAATAGATGAGAACACTCTGG 972
DB |||||
729 TGAAGATGACAAGTATATCTTTCTCGGTGAAAATGCAATAGATGAGAACACTCTGG 788
```

```
QY 973 AAAAGCTACTCAGCTAGAAATAGGTAGATATGCAAGAAATGACTTTTGGAGGGCACAGAAG 1032
DB |||||
789 AAAAGCTACTCAGCTAGAAATAGGTAGATATGCAAGAAATGACTTTTGGAGGGCACAGAAG 848
QY 1033 TCTGTGTAATAATGGACAACTTCTCTCAAGCTGCTGTGATTTCTCAGTCCAGGTCC 1092
DB |||||
849 TCTGTGTAATAATGGACAACTTCTCTCAAGCTGCTGTGATTTCTCAGTCCAGGTCC 908
QY 1093 AAATGGCAATTCACACTCATTTTGTAGAACTGCAAGATGATTTCTTAATGAATTTAAAGA 1152
DB |||||
909 AAATGGCAATTCACACTCATTTTGTAGAACTGCAAGATGATTTCTTAATGAATTTAAAGA 968
QY 1153 TCCTAAAAATCCAGTTGATATGAGAGTGTTTACAGCTTCCAGTAACTTTTCAAGGATC 1212
DB |||||
969 TCCTAAAAATCCAGTTGATATGAGAGTGTTTACAGCTTCCAGTAACTTTTCAAGGATC 1028
QY 1213 AGCCGTGTGATATAGCATGATGATGAGAGGGTGTCTTGGTCCATATATGCCCA 1272
DB |||||
1029 AGCCGTGTGATATAGCATGATGATGAGAGGGTGTCTTGGTCCATATATGCCCA 1088
QY 1273 CAGGATGGAACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCC 1332
DB |||||
1089 CAGGATGGAACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCC 1148
QY 1333 AGGAACTTTGTCCCAGCAAAACATTTTGTGGTGTGTTGACTCTACAAAGGACCTTCTCTGATGA 1392
DB |||||
1149 AGGAACTTTGTCCCAGCAAAACATTTTGTGGTGTGTTGACTCTACAAAGGACCTTCTCTGATGA 1208
QY 1393 TGTATAACCTTTTGGCAAGATCATCCAGCCATGTACAACTCAGTGTGTTTCTATGAACAA 1452
DB |||||
1209 TGTATAACCTTTTGGCAAGATCATCCAGCCATGTACAACTCAGTGTGTTTCTATGAACAA 1268
QY 1453 TCGCCCAATAGTGTACAAACGGATGTAATTTATCAATTTACAAATTTGCTGTAGACCG 1512
DB |||||
1269 TCGCCCAATAGTGTACAAACGGATGTAATTTATCAATTTACAAATTTGCTGTAGACCG 1328
QY 1513 AGTGGATGCAAGAATGGACATGATGTTTATCGGAAACAGATGTTGGAGCCGT 1572
DB |||||
1329 AGTGGATGCAAGAATGGACATGATGTTTATCGGAAACAGATGTTGGAGCCGT 1388
QY 1573 TCTTAAAGTAGTTTCAATCTCTAAGGACATGTTGATGATTTAGAAGGTTCTGCTGGA 1632
DB |||||
1389 TCTTAAAGTAGTTTCAATCTCTAAGGACATGTTGATGATTTAGAAGGTTCTGCTGGA 1448
QY 1633 AGAATGACAGTCTTTCGGGAACGACGTCTATTTACAGCAATGAGCTTTCCACTTAAGCA 1692
DB |||||
1449 AGAATGACAGTCTTTCGGGAACGACGTCTATTTACAGCAATGAGCTTTTCCACTTAAGCA 1508
QY 1693 GCAACAACTATATATTTGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTA 1752
DB |||||
1509 GCAACAACTATATATTTGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTA 1568
QY 1753 TATTTACGGGAACGGTGTGCTGATGTTGCTCGCCGAGACCTTACTGTGTTGGGA 1812
DB |||||
1569 TATTTACGGGAACGGTGTGCTGATGTTGCTCGCCGAGACCTTACTGTGTTGGGA 1628
QY 1813 TGGTTCTGCAATGTTCTCGCTATTTTCCACTGCAAGAGAGCGACAGACAGATAT 1872
DB |||||
1629 TGGTTCTGCAATGTTCTCGCTATTTTCCACTGCAAGAGAGCGACAGACAGATAT 1688
QY 1873 AAGAAATGGAGACCCACTGACTCTGCTTACAGCTTACACCATGATTAATCAACCATGGCA 1932
DB |||||
1689 AAGAAATGGAGACCCACTGACTCTGCTTACAGCTTACACCATGATTAATCAACCATGGCA 1748
QY 1933 CAGCCCTTGAAGAGAGATTCATATGTTGTAGAAATAGTAGCAATTTTGGAAATGCAG 1992
DB |||||
1749 CAGCCCTTGAAGAGAGATTCATATGTTGTAGAAATAGTAGCAATTTTGGAAATGCAG 1808
QY 1993 TCCCAAGCTCCGAGAGAGGCTGCTTATTTGCAATTTCCAGGGCAATAGAGAGCGAAA 2052
DB |||||
1809 TCCCAAGCTCCGAGAGAGGCTGCTTATTTGCAATTTCCAGGGCAATAGAGAGCGAAA 1868
QY 2053 AGAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTCTTCTGCTACGTAG 2112
```

Db	1869	AGAAGAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAG	1928
Qy	2113	TCTACACAGAGGATCAGGCAATTAACCTCTCCATCGGTGGACATGGGTTCATACA	2172
Db	1929	TCTACACAGAGGATTCAGGCAATTAACCTCTCCATCGGTGGACATGGGTTCATACA	1988
Qy	2173	AACCTCTTCTAAGGTAACTGAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCA	2232
Db	1989	AACCTCTTCTAAGGTAACTGAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCA	2048
Qy	2233	TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATGACATGACACCTAG	2292
Db	2049	TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATGACATGACACCTAG	2108
Qy	2293	CCAGAGGTCTGTGACAGAGACTTCATGACGCTCATCAACACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAGGTCTGTGACAGAGACTTCATGACGCTCATCAACACCCCAATCTCAACAGAT	2168
Qy	2353	GGATGAGTCTGTGACAAAGTTTGGAAAGGGACCGAAACCAACGCTCGCAAGGCCAGG	2412
Db	2169	GGATGAGTCTGTGACAAAGTTTGGAAAGGGACCGAAACCAACGCTCGCAAGGCCAGG	2228
Qy	2413	ACATACCCAGGGAACAGTAACAAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAA	2472
Db	2229	ACATACCCAGGGAACAGTAACAAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAA	2288
Qy	2473	CAGGAGGCCAGCAATTTGAGAGGCCACCGAGAGTGTGAGCTGCAATACCTCTAGA	2532
Db	2289	CAGGAGGCCAGCAATTTGAGAGGCCACCGAGAGTGTGAGCTGCAATACCTCTAGA	2348
Qy	2533	AACCTCAACCAAGTACAACTTGCCTAGACAATACTGGAACCAACCAATGCAATATACAT	2592
Db	2349	AACCTCAACCAAGTACAACTTGCCTAGACAATACTGGAACCAACCAATGCAATATACAT	2408
Qy	2593	GAACTTTTTCATGCGCATTTATGCGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCA	2652
Db	2409	GAACTTTTTCATGCGCATTTATGCGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCA	2468
Qy	2653	CCAAATTAATAATCAATGATGATTAACCTTCTTAATAGGCTTTTTTTC	2702
Db	2469	CCAAATTAATAATCAATGATGATTAACCTTCTTAATAGGCTTTTTTTC	2518
RESULT 4			
HUMSEM			
LOCUS			
DEFINITION		2530 bp mRNA linear	PRI 08-MAY-1995
ACCESSION		Homo sapiens semaphorin-III (Hsma-I) mRNA, complete cds.	
VERSION		L26081	
KEYWORDS		L26081.1 GI:799328	
SOURCE		semaphorin	
ORGANISM		Homo sapiens (human)	
REFERENCE			
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
JOURNAL		1 (bases 1 to 2530)	
MEDLINE		Kolodkin, A.L., Matthes, D.J. and Goodman, C.S.	
PUBMED		The semaphorin genes encode a family of transmembrane and secreted	
COMMENT		growth cone guidance molecules	
FEATURES		Cell 75 (7), 1389-1399 (1993)	
source		94094332	
Location/Qualifiers		8269517	
1..2530		On May 8, 1995 this sequence version replaced gi:436559.	
/organism="Homo sapiens"		Original source text: Homo sapiens (human).	
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="brain"			
/dev_stage="foetus"			
1..2530			
/gene="Hsma-III"			
gene			

CDS			
16..2331			
/gene="Hsma-III"			
/codon_start=1			
/product="semaphorin-III"			
/protein_id="AAA65938.1"			
/db_xref="GI:436560"			
/translation="MGWLTIRVCLFVWGLLTARANYQNGKNVPRPKLSYKEMLESNN			
VITFNGLANSSSYHTFLDLERSLYGAKDHI PFDLVNI KDFKI VWPVSYVTRRDE			
CKMAGKDLKSCANFELVLYKAYNOTHLVACGTGPHPLCTVIEIGHHPEDNIFKLENS			
HFENGRGKSPDPKLLFASLLIDELYSGTAAADPMGRDFAIFRILGHHHPITEQHDS			
RWLNDPKFISAHLLISESDNDEDDKVIYFFRENAIDGHSKATHARIIGQIKNDPFGH			
RSLVNKMWTFILKARLISVPGPNGLIDTDFDELQVFLMNFDPKPNVPYVFTSSNI			
FKGSAVCMSYMSDVRVFLGPAHYAGDNGYQVYQGRVPYPRGCTCSKTFGGDST			
KDLDDVTIIFARSHPNVPVPMNNRPIVITKDVNYQTOIVVDRVDAEDQYDMVF			
IGTDVGTIVKVSIPKETWYDLESLLEEMTVREPTAISAMELSTKQQLVIGSTAG			
VAQPLHRCDIYGKACAECLLARDPYCAWDGSACSRYEPTAKRRRRQDIRNGDPLTH			
CSDLHDHNNHHSPEERIIYGVNSSTFLECSPKSQRALVITWOPORNEERKEERVD			
DHIIIRTDGGLRLSLQKQSGNYLCHAVHGFIQTLKLVLEVIDTEHLEELHKKDD			
GDGSKTKEMSNSMTPSQKMYRDFMQLINHPNLMTDEFCEQVWVKRDRKQRQRFPQT			
PGNSNKKWHLQENKKGNRRRTHEFERAPRSV"			
BASE COUNT	786 a	518 c	576 g 650 t
ORIGIN			
Query Match	92.6%	Score 2508.4	DB 9; Length 2530;
Best Local Similarity	100.0%	Pred. No. 0;	
Matches 2509; Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
Qy	193	CTGACGATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTTACAGC	252
Db	9	CTGCAGCATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTTACAGC	68
Qy	253	AAGACAACTATCAGATGGAGAAACAATGTGCCAAGGCTGAAATTTATCTTACAAAGA	312
Db	69	AAGACAACTATCAGATGGAGAAACAATGTGCCAAGGCTGAAATTTATCTTACAAAGA	128
Qy	313	AATGTTGGAATCCAACTATGATCATTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	372
Db	129	AATGTTGGAATCCAACTATGATCATTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	188
Qy	373	TACCTTCCTTTTGGATGAGGAAACGAGTAGCTGTATGTTGGAGCAAGGATCACATAT	432
Db	189	TACCTTCCTTTTGGATGAGGAAACGAGTAGCTGTATGTTGGAGCAAGGATCACATAT	248
Qy	433	TTTCATTCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC	492
Db	249	TTTCATTCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC	308
Qy	493	CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCGAAAGATGTGCTTAATTCAT	552
Db	309	CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCGAAAGATGTGCTTAATTCAT	368
Qy	553	CAAGGTACTTAAGGCATATAATCAGATCATTGACCTGTGACGCTGTGGAACGGGGCTTTCA	612
Db	369	CAAGGTACTTAAGGCATATAATCAGATCATTGACCTGTGACGCTGTGGAACGGGGCTTTCA	428
Qy	613	TCCAAATTTGACCTACATTGAAATTTGGACATCATCTCCTGAGGACAAATATTTTAAAGCTGA	672
Db	429	TCCAAATTTGACCTACATTGAAATTTGGACATCATCTCCTGAGGACAAATATTTTAAAGCTGA	488
Qy	673	GAACTCACATTTTGAAAAACGGCCGTGGGAAGAGTCCATATACCTTAAGCTGCTGACAGC	732
Db	489	GAACTCACATTTTGAAAAACGGCCGTGGGAAGAGTCCATATACCTTAAGCTGCTGACAGC	548
Qy	733	ATCCCTTTTATAGATGAGAAATTTATCTCTGGAACCTGACCTGATTTTATGGGGCGAGA	792
Db	549	ATCCCTTTTATAGATGAGAAATTTATCTCTGGAACCTGACCTGATTTTATGGGGCGAGA	608
Qy	793	CTTTGCTATCTTCCGAACCTCTTTGGCACCCACCCCAATCAGGACAGACAGCATGATTC	852
Db	609	CTTTGCTATCTTCCGAACCTCTTTGGCACCCACCCCAATCAGGACAGACAGCATGATTC	668
Qy	853	CAGGTGGCTCAATGATCCAAAGTTTCAATTTAGTGGCCACCTCATCTCAGAGAGTGCAATCC	912

669	Qy	669	CAGTGGCTCAATGATCAAAAGTTCAATTAGTGCCCACTTATGTCGACCACTCATCTCAGAGAGTGACAAATCC	728
913	Qy	913	TGAAGATGACAAAGTATACCTTTTCTCCGTGAAAAATGCAATAGATCGAGAAACAATCTGG	972
729	Db	729	TGAAGATGACAAAGTATACCTTTTCTCCGTGAAAAATGCAATAGATCGAGAAACAATCTGG	788
973	Qy	973	AAAAGCTACTCAGCTAGAAATAGGTCAGATATGCAAGAATGACTTTTGGAGGGGCACAGAAG	1032
789	Db	789	AAAAGCTACTCAGCTAGAAATAGGTCAGATATGCAAGAATGACTTTTGGAGGGGCACAGAAG	848
1033	Qy	1033	TCTGGTGAATAAATGGACACATTCCTCAAGGCTCGTCTGATTTGGCTCAGTGCAGGTCC	1092
849	Db	849	TCTGGTGAATAAATGGACACATTCCTCAAGGCTCGTCTGATTTGGCTCAGTGCAGGTCC	908
1093	Qy	1093	AAATGGCAITGCACACTCATTTTGATGAACTGCAAGATGTATTCCTAATGAACATTTAAAGA	1152
909	Db	909	AAATGGCAITGCACACTCATTTTGATGAACTGCAAGATGTATTCCTAATGAACATTTAAAGA	968
1153	Qy	1153	TCCTAAAAATCCAGTTGTATATGAGAGTGTTTACGACTTCAGTAACAATTTTCAAGGGATC	1212
969	Db	969	TCCTAAAAATCCAGTTGTATATGAGAGTGTTTACGACTTCAGTAACAATTTTCAAGGGATC	1028
1213	Qy	1213	AGCCGTGTGTATGTATAGCATGAGTGATGTGAGNAGGGTGTTCCTTGGTCCATATGCCCA	1272
1029	Db	1029	AGCCGTGTGTATGTATAGCATGAGTGATGTGAGNAGGGTGTTCCTTGGTCCATATGCCCA	1088
1273	Qy	1273	CAGGGATGCAACCAACTATCAATGGGTGCCTTATCAAGGAAGAGTGCCTTATCCAGGCC	1332
1089	Db	1089	CAGGGATGCAACCAACTATCAATGGGTGCCTTATCAAGGAAGAGTGCCTTATCCAGGCC	1148
1333	Qy	1333	AGGAACCTGTGCCAGCAAAAACATTTGGTGGTTTGACTCTCAAGAGCACTTCTCTGATGA	1392
1149	Db	1149	AGGAACCTGTGCCAGCAAAAACATTTGGTGGTTTGACTCTCAAGAGCACTTCTCTGATGA	1208
1393	Qy	1393	TGTTATAACTTTTGCAGAGAAGTCATCAGCCATGTACAACTCAGTGTTCCTATGAACAA	1452
1209	Db	1209	TGTTATAACTTTTGCAGAGAAGTCATCAGCCATGTACAACTCAGTGTTCCTATGAACAA	1268
1453	Qy	1453	TCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAATTCGCTAGACCG	1512
1269	Db	1269	TCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAATTCGCTAGACCG	1328
1513	Qy	1513	AGTGGATGCAGAGAATGGACAGTATGATGTTATGTTTATCGGACACAGATTTGGGACCGT	1572
1329	Db	1329	AGTGGATGCAGAGAATGGACAGTATGATGTTATGTTTATCGGACACAGATTTGGGACCGT	1388
1573	Qy	1573	TCCTAAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAAGAGAGTTCTGCTGGA	1632
1389	Db	1389	TCCTAAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAAGAGAGTTCTGCTGGA	1448
1633	Qy	1633	AGAAATCACAGTTTTTCGGGAAACCGACTGTTATTTACAGCAATCGAGCTTCCACTAAGCA	1692
1449	Db	1449	AGAAATCACAGTTTTTCGGGAAACCGACTGTTATTTACAGCAATCGAGCTTCCACTAAGCA	1508
1693	Qy	1693	GCAACCACTATATATGGTTCAACGGCTGGGGTTGCCACGCTCCCTTTTACCCGGTGTGA	1752
1509	Db	1509	GCAACCACTATATATGGTTCAACGGCTGGGGTTGCCACGCTCCCTTTTACCCGGTGTGA	1568
1753	Qy	1753	TATTTACGGGAAAGCGTGTCTGAGTGTGCTCGCCCGGAGACCTTTACTGTGCTGGGA	1812
1569	Db	1569	TATTTACGGGAAAGCGTGTCTGAGTGTGCTCGCCCGGAGACCTTTACTGTGCTGGGA	1628
1813	Qy	1813	TGGTTTCGATGTTCTCGCTATTTTCCACTGCAAAAGAGACGCAACAGCAAGATAT	1872
1629	Db	1629	TGGTTTCGATGTTCTCGCTATTTTCCACTGCAAAAGAGACGCAACAGCAAGATAT	1688
1873	Qy	1873	AAGAAATGGAGACCACTGACTCTACTGTTTCAGACTTTACACCATGATAATCACCATGGCCA	1932
1689	Db	1689	AAGAAATGGAGACCACTGACTCTACTGTTTCAGACTTTACACCATGATAATCACCATGGCCA	1748
1933	Qy	1933	CAGCCCTGAAAGAGAAATCAATCTATGGTGTAGAGAAATAGTAGCACATTTTGGAAATGCAG	1992
1749	Db	1749	CAGCCCTGAAAGAGAAATCAATCTATGGTGTAGAGAAATAGTAGCACATTTTGGAAATGCAG	1808

Qy	1993	TCCGAAGTCG	CAGAGACCG	TGCTGCTATT	TGGCAATTC	CAGAGCG	GAATG	GAAGAGCG	AAA	2052
Db	1809	TCCGAAGTCG	CAGAGACCG	TGCTGCTATT	TGGCAATTC	CAGAGCG	GAATG	GAAGAGCG	AAA	1868
Qy	2053	AGAAGAGATC	AGAGTGG	ATGATCAT	ATATCAT	CAGAGAC	GAGATCA	AGGCTTC	TGCTAC	2112
Db	1869	AGAAGAGATC	AGAGTGG	ATGATCAT	ATATCAT	CAGAGAC	GAGATCA	AGGCTTC	TGCTAC	1928
Qy	2113	TCTACAAC	GAAGGATTC	CAGGCAATTA	CTCTCT	GCCATG	CGGTGG	AACATGG	GTTCAT	2172
Db	1929	TCTACAAC	GAAGGATTC	CAGGCAATTA	CTCTCT	GCCATG	CGGTGG	AACATGG	GTTCAT	1988
Qy	2173	AACCTCTCTT	TAAAGTAA	CCCTGGA	GTCATTG	TACACAG	AGCATTT	CGGAAG	AACCTCTT	2232
Db	1989	AACCTCTCTT	TAAAGTAA	CCCTGGA	GTCATTG	TACACAG	AGCATTT	CGGAAG	AACCTCTT	2048
Qy	2233	TAAAGATG	ATGATGG	AGATGG	CTTAAG	ACCAAG	AAATGT	TCCAAT	GACATGA	2292
Db	2049	TAAAGATG	ATGATGG	AGATGG	CTTAAG	ACCAAG	AAATGT	TCCAAT	GACATGA	2108
Qy	2293	CCAGAAGG	CTGGTAC	AGAGACTT	CATG	CAGCTCAT	CAAC	CCCAATCT	CAACAG	2352
Db	2109	CCAGAAGG	CTGGTAC	AGAGACTT	CATG	CAGCTCAT	CAAC	CCCAATCT	CAACAG	2168
Qy	2353	GGATGAGT	CTCTG	TCAACAA	AGTTTGG	AAAAAG	GGACCG	GAAC	CGTCGG	2412
Db	2169	GGATGAGT	CTCTG	TCAACAA	AGTTTGG	AAAAAG	GGACCG	GAAC	CGTCGG	2228
Qy	2413	ACATACCC	CGGGAA	CAGTAA	CAAAAT	TGGA	AGCATTA	CAAG	AAANAAT	2472
Db	2229	ACATACCC	CGGGAA	CAGTAA	CAAAAT	TGGA	AGCATTA	CAAG	AAANAAT	2288
Qy	2473	CAGGAGG	ACCCAG	AATTTG	AGAGGG	GCACCC	AGGAGTGT	CTG	AGCTGC	2532
Db	2289	CAGGAGG	ACCCAG	AATTTG	AGAGGG	GCACCC	AGGAGTGT	CTG	AGCTGC	2348
Qy	2533	AACCTCAA	CAAGTAG	AACCTT	GCCTAG	ACAATP	AACTGG	AAAA	CAAAATG	2592
Db	2349	AACCTCAA	CAAGTAG	AACCTT	GCCTAG	ACAATP	AACTGG	AAAA	CAAAATG	2408
Qy	2593	GAACTTTTT	TATG	GCATATG	TGGATG	TTTCA	AAATGG	TGGG	AAAA	2652
Db	2409	GAACTTTTT	TATG	GCATATG	TGGATG	TTTCA	AAATGG	TGGG	AAAA	2468
Qy	2653	CCAATTATA	AAATTA	ATCC	ATGAGT	AACTT	CTC	TAATAG	CTTTTT	2702
Db	2469	CCAATTATA	AAATTA	ATCC	ATGAGT	AACTT	CTC	TAATAG	CTTTTT	2518

RESULT 5	
G31703	
LOCUS	2530 bp DNA linear STS 28-SEP-1998
DEFINITION	sWSS1973 Eric D. Green Homo sapiens STS genomic, sequence tagged site.
ACCESSION	G31703
VERSION	G31703.1 GI:1916428
KEYWORDS	STS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2530) Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F., Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S., Leckie,W.P. and Green,E.D. A collection of 1814 human chromosome 7-specific STSs Genome Res. 7 (1), 59-64 (1997)
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	97189344
PUBMED	9037602
REFERENCE	2 (bases 1 to 2530)
AUTHORS	Green,E.D.
TITLE	Human chromosome 7 STSs (1997)

JOURNAL
COMMENT

Unpublished (1997)
 Synonyms: SEMIII
 GDB_DSEG: SEMIII
 Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735
 Email: egreen@nhgri.nih.gov
 Primer A: CCCTGGAAGTCAATGACACAG
 Primer B: GGTATGCTCGCCCTTTGCCG
 STS size: 230
 PCR Profile:
 Denaturation: 0 degrees C for 0.00 minute(s)
 Annealing: 92 degrees C for 0.17 minute(s)
 Polymerization: 55 degrees C for 1.00 minute(s)
 PCR Cycles: 35
 Thermal Cycler: PerkinElmer 9600
 Protocol:
 Template: 30-100 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul
 Buffer:
 MgCl2: 1.5 mM
 KCl: 100 mM
 Tris-HCl: 10 mM
 NH4Cl: 5 mM
 pH: 8.6

This STS has been incorporated into the NHGRI chromosome 7 physical map, but was developed by another investigator. See GenBank record: L26081 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GRS/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES

source
 1..2530
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="7"
 /clone_lib="Eric D. Green"
 gene
 1..2530
 /gene="SEMIII"
 STS
 2006..2235
 /gene="SEMIII"
 primer_bind
 2006..2026
 /gene="SEMIII"
 primer_bind
 complement(2215..2235)
 BASE_COUNT 786 a 518 c 576 g 650 t
 ORIGIN
 Query Match 92.6%; Score 2508.4; DB 11; Length 2530;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 193 CTGACGATGGGCTGTTAACTAGGATGCTGTCTTTCTGGGAGTATTACTTACAGC 252
 Db 9 CTGACGATGGGCTGTTAACTAGGATGCTGTCTTTCTGGGAGTATTACTTACAGC 68
 QY 253 AAGAGCAAACTATCAGATGGGAAGAAACAATGTGCCAGGCTGAAATTATCTCAAGA 312
 Db 69 AAGAGCAAACTATCAGATGGGAAGAAACAATGTGCCAGGCTGAAATTATCTCAAGA 128
 QY 313 AATGTGGAAATCCAAAGTGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 372
 Db 129 AATGTGGAAATCCAAAGTGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 188

QY 373 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGAGCAAAAGGATCACATATT 432
 Db 189 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGAGCAAAAGGATCACATATT 248
 QY 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTTGGCCAGATATCTTACAC 492
 Db 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTTGGCCAGATATCTTACAC 308
 QY 493 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGCTGTAATTTTCAT 552
 Db 309 CAGAAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGCTGTAATTTTCAT 368
 QY 553 CAAGGTACTTAAGGCATATAATCAGACTCACTTGTACGCCTGTGAAACGGGGCTTTTCA 612
 Db 369 CAAGGTACTTAAGGCATATAATCAGACTCACTTGTACGCCTGTGAAACGGGGCTTTTCA 428
 QY 613 TCCAATTTGCACTTACATTTGAAATTTGACATCATCTCTGAGGACATATTTTAAAGCTGGA 672
 Db 429 TCCAATTTGCACTTACATTTGAAATTTGACATCATCTCTGAGGACATATTTTAAAGCTGGA 488
 QY 673 GAACCTCAATTTTGAACACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
 Db 489 GAACCTCAATTTTGAACACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
 QY 733 ATCCCTTTTAATAGATGGAGAAATTATCTCTGGAATCTGACGTGATTTTATGGGGCGAGA 792
 Db 549 ATCCCTTTTAATAGATGGAGAAATTATCTCTGGAATCTGACGTGATTTTATGGGGCGAGA 608
 QY 793 CTTTGTATCTTCGAACTCTTGGGCACCAACCCCAATCAGGACAGAGCATGATTC 852
 Db 609 CTTTGTATCTTCGAACTCTTGGGCACCAACCCCAATCAGGACAGAGCATGATTC 668
 QY 853 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCACCTCATCTCAGAGAGTGACAATCC 912
 Db 669 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCACCTCATCTCAGAGAGTGACAATCC 728
 QY 913 TGAAGATGACAAAGTATATCTTTCTCCGTGAAAATGCAATAGATGAGGAACACTCTGG 972
 Db 729 TGAAGATGACAAAGTATATCTTTCTCCGTGAAAATGCAATAGATGAGGAACACTCTGG 788
 QY 973 AAAAGCTACTCACGCTAGATATGAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAG 1032
 Db 789 AAAAGCTACTCACGCTAGATATGAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAG 848
 QY 1033 TCTGTGTAATAAATGGCAACAATCTCCAAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC 1092
 Db 849 TCTGTGTAATAAATGGCAACAATCTCCAAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC 908
 QY 1093 AAATGGCATTCACACTCATTTTGTATGAACTGCAGGATGTAATTCCTTAATGAACTTTAAAGA 1152
 Db 909 AAATGGCATTCACACTCATTTTGTATGAACTGCAGGATGTAATTCCTTAATGAACTTTAAAGA 968
 QY 1153 TCCTAAATAATCCAGTTGTATATGAGGTTTACGACTTCCAGTTCAGTAACATTTTCAAGGGATC 1212
 Db 969 TCCTAAATAATCCAGTTGTATATGAGGTTTACGACTTCCAGTTCAGTAACATTTTCAAGGGATC 1028
 QY 1213 AGCGTGTATGTATAGCATGAGTGTGTCAGAAAGGTTTCTTGTGTCATATGCCCA 1272
 Db 1029 AGCGTGTATGTATAGCATGAGTGTGTCAGAAAGGTTTCTTGTGTCATATGCCCA 1088
 QY 1273 CAGGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCC 1332
 Db 1089 CAGGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCC 1148
 QY 1333 AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGTGACTCTCAAGGACCTTCTGTATGA 1392
 Db 1149 AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGTGACTCTCAAGGACCTTCTGTATGA 1208
 QY 1393 TGTATTAACCTTTCGAAGAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA 1452
 Db 1209 TGTATTAACCTTTCGAAGAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA 1268
 QY 1453 TCGCCCAATAGTGATCAAAACGGATGTAATTAATCAATTTTACAAATTTGCTAGACCG 1512

Db	2349	AACTCAAAACAAGTAGAAACTTGCTTAGACAATAAATCTGGAAAAACAATAATCAATATACAT	2400
Qy	2593	GAACCTTTTTCATGGCATATATGTGGATGTTTACAATGTGGGAAATTCAGCTGAGTTTCCA	2652
Db	2409	GAACCTTTTTCATGGCATATATGTGGATGTTTACAATGTGGGAAATTCAGCTGAGTTTCCA	2468
Qy	2653	CCAATTATAAATTAAATCCATGAGTAACCTTCTTAATAGGCTTTTTTTTCC	2702
Db	2469	CCAATTATAAATTAAATCCATGAGTAACCTTCTTAATAGGCTTTTTTTTCC	2518
RESULT 6			
AR040598			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2509; Conservative			
0; Mismatches			
1; Indels			
0; Gaps			
Qy	193	CTGCAGCATGGCTGGTTAACTAGGATGTCTCTTTCTGGGAGTATTACTTACAGC	252
Db	9	CTGCAGCATGGCTGGTTAACTAGGATGTCTCTTTCTGGGAGTATTACTTACAGC	68
Qy	253	AAGAGCAACTATCAGNATGGGAGAACATGTGCCAAGGCTGNAATTATCCTACAAAGA	312
Db	69	AAGAGCAACTATCAGNATGGGAGAACATGTGCCAAGGCTGNAATTATCCTACAAAGA	128
Qy	313	AATGTTGGAAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA	372
Db	129	AATGTTGGAAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA	188
Qy	373	TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT	432
Db	189	TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT	248
Qy	433	TTCAATTCGACTGGTTAAATCAAGGATTTTCAAAGATTTGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATTCGACTGGTTAAATCAAGGATTTTCAAAGATTTGTGTGGCCAGTATCTTACAC	308
Qy	493	CAGAAGAGATGAATGCAAGTGGGCTGGAAAGAATCTCTGAAAGAAATGTCTAAATTTTCAT	552
Db	309	CAGAAGAGATGAATGCAAGTGGGCTGGAAAGAATCTCTGAAAGAAATGTCTAAATTTTCAT	368
Qy	553	CAAGGTACTTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGGCTTTTCA	612
Db	369	CAAGGTACTTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGGCTTTTCA	428
Qy	613	TCCAAATTGCACTCAATTTGAAATTTGGAACATCATCTCTGAGGACAATATTTTTTAAGCTGGA	672
Db	429	TCCAAATTGCACTCAATTTGAAATTTGGAACATCATCTCTGAGGACAATATTTTTTAAGCTGGA	488
Qy	673	GAACCTCAATTTTGAAACCGGCGTGGAGAGTCCCATATGACCCCTAAGCTGTGACAGC	732
Db	489	GAACCTCAATTTTGAAACCGGCGTGGAGAGTCCCATATGACCCCTAAGCTGTGACAGC	548
Qy	733	ATCCCTTTTAAATAGATGGAGAAATTATACTCTGGAACTGCAGCTGATTTTTATGGGGGAGA	792

|||||
549 ATCCCTTTTAAATAGATGGAGAAATTACTCTGGAACTGCAGCTGATTTTATGGGGGAGA 608
|||||
793 CTTTGTCTATCTTCGGAACCTTTGGGACACCAACCACCAATCAGGACAGACGATGATTC 852
|||||
609 CTTTGTCTATCTTCGGAACCTTTGGGACACCAACCACCAATCAGGACAGACGATGATTC 668
|||||
853 CAGGTGGCTCAATGATCCAAAGTTCAATTAGTGCCACCTCATCTCAGAGAGTGACAATCC 912
|||||
669 CAGGTGGCTCAATGATCCAAAGTTCAATTAGTGCCACCTCATCTCAGAGAGTGACAATCC 728
|||||
913 TGAAGATGACAAAGTATACCTTTTCTTCGCTGGAATGCAATAGATGAGAACTCTCG 972
|||||
729 TGAAGATGACAAAGTATACCTTTTCTTCGCTGGAATGCAATAGATGAGAACTCTCG 788
|||||
973 AAAAGTACTCAGCTAGATAGGTGATGATGCAAGATGATGTTGGAGGGGACAGAAG 1032
|||||
789 AAAAGTACTCAGCTAGATAGGTGATGATGCAAGATGATGTTGGAGGGGACAGAAG 848
|||||
1033 TCTGTGTAATAATGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 1092
|||||
849 TCTGTGTAATAATGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 908
|||||
1093 AAATGSCATTGACACTCATTTTGTATGAACCTGCAGGATGATTCCTAATGAACCTTTAAAGA 1152
|||||
909 AAATGSCATTGACACTCATTTTGTATGAACCTGCAGGATGATTCCTAATGAACCTTTAAAGA 968
|||||
1153 TCTTAAATAATCAGTTGTATATGAGAGTTTACGACTTCCAGTAACTTTTCAAGGGATC 1212
|||||
969 TCTTAAATAATCAGTTGTATATGAGAGTTTACGACTTCCAGTAACTTTTCAAGGGATC 1028
|||||
1213 AGCCGTGCTGATCTATAGCATGATGATGTGAGAGGGTTCCTTGGTCCATATGCCCCA 1272
|||||
1029 AGCCGTGCTGATCTATAGCATGATGATGTGAGAGGGTTCCTTGGTCCATATGCCCCA 1088
|||||
1273 CAGGGATGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGGCC 1332
|||||
1089 CAGGGATGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGGCC 1148
|||||
1333 AGGAACCTTGCAGCAAAACATTTGGTGGTTCGATCTACAAAGGACCTTCCTGATGA 1392
|||||
1149 AGGAACCTTGCAGCAAAACATTTGGTGGTTCGATCTACAAAGGACCTTCCTGATGA 1208
|||||
1393 TGTATAAATCTTGCAGCAAAACATTTGGTGGTTCGATCTACAAAGGACCTTCCTGATGA 1452
|||||
1209 TGTATAAATCTTGCAGCAAAACATTTGGTGGTTCGATCTACAAAGGACCTTCCTGATGA 1268
|||||
1453 TCGCCCAATAGTATCAAAACCGGATGATTAATTTATCAAAATTTGCTGAGACCG 1512
|||||
1269 TCGCCCAATAGTATCAAAACCGGATGATTAATTTATCAAAATTTGCTGAGACCG 1328
|||||
1513 AGTGGATGACAGAGATGGACAGTATGATGTTATTCGGAAACAGATGTTGGACCGT 1572
|||||
1329 AGTGGATGACAGAGATGGACAGTATGATGTTATTCGGAAACAGATGTTGGACCGT 1388
|||||
1573 TCTTAAAGTATGTTTCAATTTCTTAAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1632
|||||
1389 TCTTAAAGTATGTTTCAATTTCTTAAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1448
|||||
1633 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1692
|||||
1449 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1508
|||||
1693 GCACAACTATATATGTTTCAACGGCTGGGTGGCCAGCTCCCTTTACACGGTGTGA 1752
|||||
1509 GCACAACTATATATGTTTCAACGGCTGGGTGGCCAGCTCCCTTTACACGGTGTGA 1568
|||||
1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGTGGA 1812
|||||
1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGTGGA 1628
|||||
1813 TGTGTTCTGATGTTCTCGTATTTTCCCACTGCTCAAGAGACGCAAGACGACAAGATAT 1872
|||||

Db Db
1629 TGGTCTTCGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGACAAGATAT 1688
Qy Qy
1873 AAGAAATGAGACCCCACTGACTCAGCTGTTTACCACTGATATATACCATGGCCA 1932
Db Db
1689 AAGAAATGAGACCCCACTGACTCAGCTGTTTACCACTGATATATACCATGGCCA 1748
Qy Qy
1933 CAGCCCTGAAGAGAGAATCATCTATGTTGTAGAGAATAGTAGACAATTTTGGAAATGCA 1992
Db Db
1749 CAGCCCTGAAGAGAGAATCATCTATGTTGTAGAGAATAGTAGACAATTTTGGAAATGCA 1808
Qy Qy
1993 TCCGAAGTCCGAGAGAGCGCTGCTTATGGAATTCAGAGGCGGAATGAAGAGCGAAA 2052
Db Db
1809 TCCGAAGTCCGAGAGAGCGCTGCTTATGGAATTCAGAGGCGGAATGAAGAGCGAAA 1868
Qy Qy
2053 AAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 2112
Db Db
1869 AAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 1928
Qy Qy
2113 TCTCAACAGAGAGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTTCATACA 2172
Db Db
1929 TCTCAACAGAGAGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTTCATACA 1988
Qy Qy
2173 AACTCTTCTTAAGGTAACCTCGGAAGTCAATGACACAGAGATTTGGAAAGAACTTCTTCA 2232
Db Db
1989 AACTCTTCTTAAGGTAACCTCGGAAGTCAATGACACAGAGATTTGGAAAGAACTTCTTCA 2048
Qy Qy
2233 TAAAGATGATGATGGAGATGCTTAAAGCAAGAAATGTCATAGCATGACACCTAG 2292
Db Db
2049 TAAAGATGATGATGGAGATGCTTAAAGCAAGAAATGTCATAGCATGACACCTAG 2108
Qy Qy
2293 CCAGAGGCTGCTGATCAGAGACTTTCATGAGCTCATCAACCAACCCCAATCTCAACACGAT 2352
Db Db
2109 CCAGAGGCTGCTGATCAGAGACTTTCATGAGCTCATCAACCAACCCCAATCTCAACACGAT 2168
Qy Qy
2353 GGATGAGTCTGTGAACAAAGTTTGAAGAGGACCGCAAAACAACTCGGCAAGGCGAG 2412
Db Db
2169 GGATGAGTCTGTGAACAAAGTTTGAAGAGGACCGCAAAACAACTCGGCAAGGCGAG 2228
Qy Qy
2413 ACATACCCCGGAGACAGTAACTGAAGACCTTACAGAAATTAAGAAAGTAGAAA 2472
Db Db
2229 ACATACCCCGGAGACAGTAACTGAAGACCTTACAGAAATTAAGAAAGTAGAAA 2288
Qy Qy
2473 CAGGAGGACCCCAATTTGAGAGGACCCAGGAGTCTGCTGAGCTGATACCTCTAGA 2532
Db Db
2289 CAGGAGGACCCCAATTTGAGAGGACCCAGGAGTCTGCTGAGCTGATACCTCTAGA 2348
Qy Qy
2533 AACCTCAAAAGTAGAACTTGGCTGACAAATTAAGTGAAGAAACAAATGCAATATACAT 2592
Db Db
2349 AACCTCAAAAGTAGAACTTGGCTGACAAATTAAGTGAAGAAACAAATGCAATATACAT 2408
Qy Qy
2593 GAACTTTTTCATGCGCATTTATGAGTGTTCATGATGTTTACATGTTGGAAATTCAGCTGAGTTCCA 2652
Db Db
2409 GAACTTTTTCATGCGCATTTATGAGTGTTCATGATGTTTACATGTTGGAAATTCAGCTGAGTTCCA 2468
Qy Qy
2653 CCAATTATAAATTAATCCATGAGTAACTTTTCCATAGGCTTTTTCCTCC 2702
Db Db
2469 CCAATTATAAATTAATCCATGAGTAACTTTTCCATAGGCTTTTTCCTCC 2518

RESULT 7

LOCUS 147054
DEFINITION Sequence 53 from patent US 5639856.
ACCESSION 147054
VERSION 147054.1 GI:2471019
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Goodman, C.S., Kolodkin, A.L., Matthes, D., Bentley, D.R. and O'Connor, T.
TITLE Semaphorin gene family

JOURNAL Patent: US 5639856-A 53 17-JUN-1997;

FEATURES Location/Qualifiers

source 1..2601

BASE COUNT 809 a 533 c 593 g 666 t

ORIGIN

Query Match 92.6%; Score 2508.4; DB 6; Length 2601;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCACATGGGCTGTTAACTAGGATTCCTGTCTTTCTGGGAGTATTACTTACAGC	252
DB	9	CTGCACATGGGCTGTTAACTAGGATTCCTGTCTTTCTGGGAGTATTACTTACAGC	68
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTCACAGGCTGAAATATCTTACAAGA	312
DB	69	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTCACAGGCTGAAATATCTTACAAGA	128
QY	313	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372
DB	129	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCCTTTGGATGAGGACGGAGTAGGCTGTATGTTGAGCAAGGATCAATATT	432
DB	189	TACCTTCCTTTGGATGAGGACGGAGTAGGCTGTATGTTGAGCAAGGATCAATATT	248
QY	433	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	308
QY	493	CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTTGAAGAATGTGCTAAATTTCA	552
DB	309	CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTTGAAGAATGTGCTAAATTTCA	368
QY	553	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGCTTTTCA	612
DB	369	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGCTTTTCA	428
QY	613	TCCAAATTCGACCTACATGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGA	672
DB	429	TCCAAATTCGACCTACATGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGA	488
QY	673	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	732
DB	489	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	548
QY	733	ATCCCTTTTAAATAGATGGAGATTTATCTCTGGAATCGAGCTGATTTTATGGGGCGAGA	792
DB	549	ATCCCTTTTAAATAGATGGAGATTTATCTCTGGAATCGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTGCTATCTTCCGAACTCTTGGGCAACCAACCAATCAGGACAGAGCAGCATGATTC	852
DB	609	CTTTGCTATCTTCCGAACTCTTGGGCAACCAACCAATCAGGACAGAGCAGCATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTTCAATAGTCCCAACCTCATCTCAGAGAGTGACAATCC	912
DB	669	CAGGTGGCTCAATGATCCAAAGTTTCAATAGTCCCAACCTCATCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATACCTTTTCTCGTGAAATGCAATAGATGGGAACACATCTGG	972
DB	729	TGAAGATGACAAAGTATACCTTTTCTCGTGAAATGCAATAGATGGGAACACATCTGG	788
QY	973	AAAAGCTACTCAGCTAGAAATAGGTCAGATATGCAAGATGACTTTGGGGGCGACAGAG	1032
DB	789	AAAAGCTACTCAGCTAGAAATAGGTCAGATATGCAAGATGACTTTGGGGGCGACAGAG	848
QY	1033	CTGTGTAATAATGACAACTATCTTCAAGCTGCTGTGATTTGCTCAGTCCAGGTCC	1092
DB	849	CTGTGTAATAATGACAACTATCTTCAAGCTGCTGTGATTTGCTCAGTCCAGGTCC	908
QY	1093	AAATGCGATGACATCTTTTGTGAACTGCGAGATGATTTCTTAATGAATCTTTAAAGA	1152

DB	909	AAATGCGCATTGACACTCAATTTTGTATGAACGAGATGTATTCCTAATGAACCTTTAAAGA	968
QY	1153	TCCTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATC	1212
DB	969	TCCTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATC	1028
QY	1213	AGCCGTGTATGTATAGCATGATGTGAGAAAGGTGTTCCTTGGTCCATATATGCCCA	1272
DB	1029	AGCCGTGTATGTATAGCATGATGTGAGAAAGGTGTTCCTTGGTCCATATATGCCCA	1088
QY	1273	CAGGATGAGCCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGAGCCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1148
QY	1333	AGGAATCTGTCCAGCAAAACATTTGGTGGTTTTCACCTCTACAAAGGACCTTCTCTATGA	1392
DB	1149	AGGAATCTGTCCAGCAAAACATTTGGTGGTTTTCACCTCTACAAAGGACCTTCTCTATGA	1208
QY	1393	TGTTATAAATCTTTTGAAGAATCATCCAGGCCATGTACAAATCCAGTGTTCCTATGAACAA	1452
DB	1209	TGTTATAAATCTTTTGAAGAATCATCCAGGCCATGTACAAATCCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCCAATAGTGTATCAAAACGGATGTAAATTTTACAAATTTTACAAATTTGCTAGACG	1512
DB	1269	TCGCCCAATAGTGTATCAAAACGGATGTAAATTTTACAAATTTTACAAATTTGCTAGACG	1328
QY	1513	AGTGATGACAGAAATGACAGATGATGTATTTATCGGAAACAGATTTGGGACCGT	1572
DB	1329	AGTGATGACAGAAATGACAGATGATGTATTTATCGGAAACAGATTTGGGACCGT	1388
QY	1573	TCTTAAAGTAGTTTCAATTCCTAAAGGACATTTGGTATGATTTAGAAAGGTTCCTCTGA	1632
DB	1389	TCTTAAAGTAGTTTCAATTCCTAAAGGACATTTGGTATGATTTAGAAAGGTTCCTCTGA	1448
QY	1633	AGAAATGACATTTTTCGGGAACCGACTGCTATTTTACCAATGAGCTTCCACTAAGCA	1692
DB	1449	AGAAATGACATTTTTCGGGAACCGACTGCTATTTTACCAATGAGCTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTA	1752
DB	1509	GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTA	1568
QY	1753	TATTTACGGGAAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1812
DB	1569	TATTTACGGGAAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1628
QY	1813	TGTTTCTGCTATGCTGCTATTTTCCCACTGCAAGAGAGAGCAGCAAGAGATAT	1872
DB	1629	TGTTTCTGCTATGCTGCTATTTTCCCACTGCAAGAGAGAGCAGCAAGAGATAT	1688
QY	1873	AAGAAATGAGACCCACTGACTCACTGTTTACAGACTTACACCATGATAATCACCAGGCA	1932
DB	1689	AAGAAATGAGACCCACTGACTCACTGTTTACAGACTTACACCATGATAATCACCAGGCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAAATAGTAGCAATTTTGGATGCG	1992
DB	1749	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAAATAGTAGCAATTTTGGATGCG	1808
QY	1993	TCCGAAGTCCGAGAGAGCGCTGCTTATTTGCAATTTCCAGAGCGGAAATGAGAGCGAAA	2052
DB	1809	TCCGAAGTCCGAGAGAGCGCTGCTTATTTGCAATTTCCAGAGCGGAAATGAGAGCGAAA	1868
QY	2053	AGAAGAGATCAGAGTGGATGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG	2112
DB	1869	AGAAGAGATCAGAGTGGATGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG	1928
QY	2113	TCTTAAACAGAGAGGATTTCAGGCAATTTACCTCTGCGATGCGGTGGAACATGGGTTTCA	2172
DB	1929	TCTTAAACAGAGAGGATTTCAGGCAATTTACCTCTGCGATGCGGTGGAACATGGGTTTCA	1988
QY	2173	AACCTCTTCTTAAAGTAACTTGGAGTCAATTTGACAGAGATTTGGAAGAACTTCTTCA	2232
DB	1989	AACCTCTTCTTAAAGTAACTTGGAGTCAATTTGACAGAGATTTGGAAGAACTTCTTCA	2048

```

QY 2233 TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCTCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCTCAATAGCATGACACCTAG 2108
QY 2293 CCAAGAGGCTGCTGTAAGAGATCTTCAATGAGCTCAATACCAAGGCTCAATACCAAGAT 2352
Db 2109 CCAAGAGGCTGCTGTAAGAGATCTTCAATGAGCTCAATACCAAGGCTCAATACCAAGAT 2168
QY 2353 GGATGAGTCTGTAAGAGATCTTCAATGAGCTCAATACCAAGGCTCAATACCAAGGCT 2412
Db 2169 GGATGAGTCTGTAAGAGATCTTCAATGAGCTCAATACCAAGGCTCAATACCAAGGCT 2228
QY 2413 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAAAGAAATAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAAAGAAATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCTAGA 2348
QY 2533 AACCTCAACAGTGAAGAACTTGCCTAGACAAATACCTGGAAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAACAGTGAAGAACTTGCCTAGACAAATACCTGGAAGAAACAAATGCAATATACAT 2408
QY 2593 GAACTTTTTCATGGCAATATGTCGATGTTTCAATGTCGGAATTTCAAGTGGTTCCTCA 2652
Db 2409 GAACTTTTTCATGGCAATATGTCGATGTTTCAATGTCGGAATTTCAAGTGGTTCCTCA 2468
QY 2653 CCAATTAATAATTAATCAATGAGTAACTTCTCAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTAATAATTAATCAATGAGTAACTTCTCAATAGGCTTTTTCCTCC 2518

```

RESULT 8

```

MUSC1
LOCUS Mus musculus sema3A mRNA for semaphorin 3A, complete cds.
DEFINITION D85028
ACCESSION D85028
VERSION 2
KEYWORDS collapse-1.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2356)
AUTHORS Taniguchi, M., Yusa, S., Fujisawa, H., Naruse, I., Saga, S., Mishina, M.
and Yagi, T.
TITLE Disruption of semaphorin III/D gene causes severe abnormality in peripheral nerve projection
JOURNAL Neuron 19 (3), 519-530 (1997)
MEDLINE 97470885
PUBMED 9331345
REFERENCE 2 (bases 1 to 5952)
AUTHORS Taniguchi, M.
JOURNAL Direct Submission
TITLE Submitted (02-MAY-1996) Masahiko Taniguchi, Department of Biochemistry and Molecular Biology, Graduate School of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan (E-mail: taniguti@m.u-tokyo.ac.jp, Tel: 81-3-5802-2925, Fax: 81-3-3813-8732)
COMMENT On Jan 28, 2002 this sequence version replaced gi:1313903.
FEATURES
Source
1..5952
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..5952
/gene="sema3A"
652..2970
/gene="sema3A"
/codon_start=1
/product="semaphorin 3A"

```

```

/protein_id="BAA19773.1"
/db_xref="GI:2055316"
/translation="MGWFTGTAFLFWGVLTLARANYANGKNVPLKLSYKEMLSNN
VITFNLGANSSTYHTFLDEERSLYGKADHIFSNLVNFKDFKIVPVSVYTRDE
CFKNGKDIKECANPIKIVLEYNQTHLYACGTGAFPHICTYIEVGHEDNIFKLQDS
HFNWGLGSPYDPKLLTASLIDGELYSGTAFMRDFAIFRTIRLGHHPDIFKQDHS
RWLNDPRFISAHLIPESDNPDQKVFYFFRENAIDGHSKATHARIQOICKNDFFGH
RSLVNWKTTFILKARLICVSPGNGIDTFHDLQVFLMNSKDPKNPIVGVFTTSSNI
FKGSACVYMSDSVRRVFLGAYHRDGNVQVYVQGRVPVPRPCTPSKTFGGFDS
KDLDDVITFARSHPMYNPVPIINRPIIMIKTDVNYQTOIVDVRDAEDQYDMF
IGTDVTVLVKVSPTWHDLEEVLEEMVIFREPTTISAMELSTKQQLYIGSTAG
VAQPLHRCDIYKACACECCARDPYCAWDGSSCSRYPTAKRRTTRDNRDGLPIRM
CSDLQHDNHRHPSLEERIIYVENSTSTFLCSPKSRALVYVQQRNEDRIRH
GDHIIIRTSQGLLSLOKXGNYLCHAVEHGMOTLLKVLTEVDTHLEHLLHKKD
DGDGSKLKEMSSMTPSOKVYRDFMOLINHPNLNMTDFECEQVWKRDKRQRQPGH
SQSSNKNWQESKKGRRNTHFERAPRSV"
BASE COUNT 1829 a 1203 c 1214 g 1706 t
Query Match 75.7%; Score 2050; DB 10; Length 5952;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2281; Conservative 0; Mismatches 295; Indels 14; Gaps 4;
QY 118 TTTCCAGATTGTTGAACTTCTCTGGCCGCAATACAGGAAGGAAGCTGCCGAGCTC 629
Db 570 TTTCCAGATTGTTGAACTTCTCTGGCCGCAATACAGGAAGGAAGCTGCCGAGCTC 629
QY 178 AGGACCTTACAGCTCTCAGCATGGCTGTAACTAGGATGTCCTCTTTCTGGG 237
Db 630 AGGACCTTACAGCTCTCAGCATGGCTGTAACTAGGATGTCCTCTTTCTGGG 689
QY 238 AGTATTACTTACAGCAAGCAAACTATCAGAAATGGGAAGAACTATGTCGCAAGCTGAA 297
Db 690 TGTATTACTTACAGCAAGCAAACTATGCAACGGAAGAACTATGTCGCAAGCTGAA 749
QY 298 ATTATCTTACAAAGAAATGTTGGAATCAAACTATGTCATCTTTCAATGGCTTGGCCAA 357
Db 750 ATTATCTTACAAAGAAATGTTGGAATCAAACTATGTCATCTTTCAATGGCTTGGCCAA 809
QY 358 CAGCTCCAGTTATCATACCTTCTTTGGATGAGGACGAGTAGGCTGATTTGGAGC 417
Db 810 CAGCTCCAGTTATCATACCTTCTTTGGATGAGGACGAGTAGGCTGATTTGGAGC 869
QY 418 AAAGATCACATATTTTTCATTCGACCTCTGTTTAAATATCAAGGATTTTCAAAAGATTTGTG 477
Db 870 AAAAGATCATATATTTTTCATTCGACCTCTGTTTAAATATCAAGGATTTTCAAAAGATTTGTG 929
QY 478 GCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGA 537
Db 930 GCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGA 989
QY 538 ATGTGCTAATTTTCATCAGGTACTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 597
Db 990 ATGTGCTAATTTTCATCAGGTACTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 1049
QY 598 AACGGGGCTTTTTCATCCAAATTTGACCTATCAATGAAATGGACATCATCTCTGAGGACAA 657
Db 1050 AACTGGGGCTTTTTCATCCAAATTTGACCTATCAATGAAATGGACATCATCTCTGAGGACAA 1109
QY 658 TATTTTAAAGCTGGAGAACTCAATTTTGAAGAACGGCGCTGGGAAGAGTCCATATGACCC 717
Db 1110 CATTTTAAAGCTGGAGAACTCAATTTTGAAGAACGGCGCTGGGAAGAGTCCATATGATCC 1169
QY 718 TAAGCTGCTGACAGATCCCTTTTAAATAGATGGAATATATCTCTGAACTGGAGCTGA 777
Db 1170 CAAACTACTGACTGCTCTCTTCTTAATAGACGGTGTGTACTCTGGAATGTGCTCGGA 1229
QY 778 TTTTATGGGGGAGACTTTTGCTATCTTCCGAATCTTTGGGACCAACCAACCAATCAGGAC 837
Db 1230 CTTTATGGGGGAGACTTTTGCTATCTTCCGAATCTTTGGGACCAACCAACCAATCAGGAC 1289
QY 838 AGACGACATGATTTCCAGGTGGCTCAATGATCCAAAGTTTCAATGATGTCCTCCACCTCATCTC 897

```

Db 1290 GGAGCAGATGACTCCCGGTGGCTCAATGATCTAGATTCATCAGTGGCCATCTCATCCC 1349
QY 898 AGAGAGTGCAATCTCTGAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGA 957
Db 1350 AGAGAGTGCAACCCCTGAAGATGACAAAGTATATCTTTTCTTCCGGAATGCAATAGA 1409
QY 958 TGGAGAACACTCTGGAAGAGCTACTCACGCTAGATAGGTAGATATGCAAGATGACTT 1017
Db 1410 CGGAGAACACTCTGGAAGAGCCACTCATGCTAGAATAGGTAGATATGCAAGATGACTT 1469
QY 1018 TGGAGGCCACAGAGTCTGCTGAATAAATGGCAACATCTCTCAAGCTCGTCTGATTG 1077
Db 1470 TGGTGACACAGAGTCTTGTGAATAAATGGCAACATCTCTCAAGCTCGTCTGATTG 1529
QY 1078 CTCAGTGCCAGGTCCAAATGGCAATGACACTCATTTTGTATGAATGCAAGTGTATTCCT 1137
Db 1530 CTCGTGCGCGTCCCAATGGCAATGACACTCATTTTGTATGAATGCAAGTGTATTCCT 1589
QY 1138 AATGAATCTTAAAGATCTTAAATTCAGTGTGTATATGAGTGTATTCAGTGTCCAGTAA 1197
Db 1590 AATGAATCTTAAAGATCTTAAATTCAGTGTGTATATGAGTGTATTCAGTGTCCAGTAA 1649
QY 1198 CATTTTCAAGGATCAGCGCTGTGTATGATAGCATGAGTGTATGAGTGTATTCAGTGTCC 1257
Db 1650 CATCTTAAAGGATCTGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATG 1709
QY 1258 TGGTCCATATGCCCAGGATGGCCAACTATCAATGGGTGCTTATCAAGGAAGAGT 1317
Db 1710 TGGTCCATATGCTCAGAGATGCTCCAACTATCAGTGGGTGCTTACCAGGAAGAGT 1769
QY 1318 CCCCTATCCACGCGCAGGAACCTTGTCCAGCAAAACATTTGTGTGTGTGTGTGTGTGTGT 1377
Db 1770 CCCCTATCCACGCGCAGGAACCTTGTCCAGTAAACATTTGTGTGTGTGTGTGTGTGTGT 1829
QY 1378 GGCCTTCTGTATGATGTTATTAACCTTTCAGAGATGATCCAGCCATGTAATCCAGT 1437
Db 1830 GGCCTTCTGTATGATGTTATTAACCTTTCAGAGATGATCCAGCCATGTAATCCAGT 1889
QY 1438 GTTCTCTATGAACATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACA 1497
Db 1890 GTTCTCTATGAATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACA 1949
QY 1498 AATGTGCTGAGACCGAGTGGATGCGAAGATGCGACAGTATGATGATGATGATGATGATGAT 1557
Db 1950 AATGTGCTGAGACCGAGTGGATGCGAAGATGCGACAGTATGATGATGATGATGATGATGAT 2009
QY 1558 AGATGCTGGAGCGGTCTTAAAGTAGTTCATTTCTAAGGAGACTTGTGTATGATTTAGA 1617
Db 2010 AGATGCTGGAGCGGTCTTAAAGTAGTTCATTTCTAAGGAGACTTGTGTATGATTTAGA 2069
QY 1618 AGAGGTTCTGCTGGAGAAATGACAGTCTTTCGGGAACCGAGTCTATTTTCAAGCAATGGA 1677
Db 2070 AGAGGTTCTGCTGGAGAAATGACAGTCTTTCGGGAACCGAGTCTATTTTCAAGCAATGGA 2129
QY 1678 GTTTCCTCAAGCAGCAACAACTATATATGTTTCAAGGCTGGGGTGGCCAGCTCCC 1737
Db 2130 GCTTTCTACTAAACAGCAACAGCTGTATCTGGCTCAACTGCGGAGTGGCAGAGCTTCC 2189
QY 1738 TTTACACCGGTGTATTTACGGGAAGCGTGTGCTAGTGTGTGCTGCGCCCGAGAGCC 1797
Db 2190 TTTACACCGGTGTATTTACGGGAAGCGTGTGCTAGTGTGTGCTGCGCCCGAGAGCC 2249
QY 1798 TTACTGTGCTGGAGTGTGCTGATGTTCTCGTATTTTCCACTGCAAGAGAGCGAC 1857
Db 2250 TTACTGTGCTGGAGTGTGCTGATGTTCTCGTATTTTCCACTGCAAGAGAGCGAC 2309
QY 1858 AAGACGACAAGATATAAGAAATGGAGACCTGACTCTACTGTTTCAAGTCTTCAAGTCTT 1914
Db 2310 AAGACGACAAGATATAAGAAATGGAGACCTGACTCTACTGTTTCAAGTCTTCAAGTCTT 2369
QY 1915 TGATATACCATGCGCCACAGCCCTGGAAGAGAGATTCATCTATGTTGTAGAGAAATAGTAG 1974
Db 2370 TGATATATCATGCGCCCGCCCTTGAAGAGAGATTCATCTATGAGTGGAAACAGTAG 2429

QY 1975 CACATTTTGGATGCACTCCGAAAGTCGACAGAGCGCTGCTATTTGGCAATTCAGAG 2034
Db 2430 TACATTTTGGATGCACTCCGAAAGTCGACAGAGCGCTTGGTATATTTGGCAATTCAGAG 2489
QY 2035 GCGAATATGAAGCGAAGAAAGAGATCAGAGTGGATGATCATATATCATCAGGACAGATCA 2094
Db 2490 GAGAATATGAAGTCGAAAGAGGAGATCAGAAATGGGTGATCATATATCATCAGGACAGAA 2549
QY 2095 AGGCCTTCTGCTAGCTAGTCTACACAAAGATTCAGGCAATTAACCTCTGCGCATGCCGT 2154
Db 2550 AGGCCTTCTGCTAGCTAGCTAGTCTGCAAGAGAGATTCAGGCAATTAACCTCTGCGCATGCCGT 2609
QY 2155 GGAACATCGGTTTCATACAAACTCTTCTTAAAGGTAAACCTTGAAGTCAATTTGACACAGACA 2214
Db 2610 GGAACATCGGTTTCATACAAACTCTTCTTAAAGGTAAACCTTGAAGTCAATTTGACACAGACA 2669
QY 2215 TTTGGAAGAACTCTTCTTAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTC 2274
Db 2670 TTTGGAAGAACTCTTCTTAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTC 2729
QY 2275 CAATAGCATGACACTAGCCAGAAAGTCTGCTGACAGAGACTTTCATGACAGCTCATCAACA 2334
Db 2730 GAGCAGCATGACAGCCAGCCAGAAAGTCTGCTGACAGAGACTTTCATGACAGCTCATCAACA 2789
QY 2335 CCCCAATCTCAACAGCATGAGTGTCTGTCAACAAGTCTTGGAAAAGGGACCGAAAACA 2394
Db 2790 CCCCAATCTCAACAGCATGAGTGTCTGTCAACAAGTCTTGGAAAAGGGACCGAAAACA 2849
QY 2395 AGCTGGCAAAAGGCGCAGACATACCCAGGGAACAGTAAACAAATGGAAGAGCTTTACAAGA 2454
Db 2850 AGCTGGCAAAAGGCGGCGCACTCTCAAGGGAGCAGCAACAAAGTGGAAAGCAGATGCAAGA 2909
QY 2455 AAATGAAGAAAGTGAACACAGAGGACCCAGAAATTTGAGAGGGCACCAGAGGTCTG 2514
Db 2910 GAGCAAGAAAGTGAACACAGAGGACCCAGAGTTTGGGGGGCACCAGAGGTCTG 2969
QY 2515 AGCTGATCTTCTTAGAAACCTCAACAAGTGAAGAACTTCGCTTAGACAAATAAAGTGGAAA 2574
Db 2970 AGCTGGCCACCTCCCAAAACCTCAACAAGTGAAGAACTTCGCTTAGACAAATAAAGTGGAAA 3029
QY 2575 AACAAATGCAATATACATGAACCTTTTTCATGGGCAATATGAGATGTTTCAATAGTGGG 2634
Db 3030 ---AAATGCAATATACATGAAC--ATTTCATGGCAATATGAGTGGATGTTTCAATAGTGGG 3084
QY 2635 AAATTCAGTCTGAGTTCCACCAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTA 2694
Db 3085 AAGTTCAACCGGTTCCACCAAT-----TTAAGTCCGTGAGGAACCTTTCCAGCAGGCT 3138
QY 2695 TTTTTCCTTA 2704
Db 3139 TTTTCAGTA 3148

RESULT 9

MMRNASEMD
LOCUS M.musculus mRNA for semaphorin D. 2913 bp mRNA linear ROD 08-JUL-1996
DEFINITION X85993
ACCESSION X85993.1 GI:854329
VERSION semaphorin; semD gene.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Puschel, A.W., Adams, R.H. and Betz, H.
TITLE Murine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension
JOURNAL Neuron 14 (5), 941-948 (1995)
MEDLINE 95267431
PUBMED 7748561
REFERENCE 2 (bases 1 to 2913)


```
QY 1618 AGAGGTTCTGCTGGAAGAAATGACAGTCTTTTCGGAACCGACTGCTATTTTCAGCAATGGA 1677
Db 1529 AGAAGTTCTTCTGGAAGAAATGACCGTCTTCGGAACCAACAACTATTTTCGCAATGGA 1588
QY 1678 GCTTTCCACTAAGCAGCACTATATATGCTTTCACGGCTGGGTGCTCCAGCTCC 1737
Db 1589 GCTTTCTACTAAACAGCAACAGCTGTACATTTGGCTCAACTGCGGAGTGCGACAGCTCC 1648
QY 1738 TTTACACCGGTGTGATTTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCTCGCCGAGAGACC 1797
Db 1649 TCTACACCGCTGTGACATCTATGGCAAGCGTGTGCAAGTGTGCTCGCTCGGAGACC 1708
QY 1798 TTACTGTGCTGGAGTGTGCTGATTTCTGCTATTTTCCACTGCAAGAGAGCGAC 1857
Db 1709 TTACTGTGCTGGAGTGTGCTGCTATGCTCAGCTATTTTCTACTGCAAGAGAGCGAC 1768
QY 1858 AAGACGACAGATATAAGAAATGGAGACCCACTGACTCACTGCTTCAGACTT---ACACCA 1914
Db 1769 AAGACGACAGATATAAGAAATGGAGACCCACTGACTCACTGCTTCAGACTTGGAGACCA 1828
QY 1915 TGATAATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAG 1974
Db 1829 TGATAATCATATGGCCAGCGCTTGAAGAGAGATCATCTATGAGTGGAGAAACAGTAG 1888
QY 1975 CACATTTTGGAAATGCAATCCGAAAGTCCGAGAGCGCTGTCTATTGGCAATCCAGAG 2034
Db 1889 TACATTTTGGAAATGCAATCCGAAAGTCCGAGAGCGCTTGTATATTTGGCAATTTTCA 1948
QY 2035 GCGAATGAAGCGCAAGAGAGATCAGAGTGGATGATCATATCATCAGCAGACAGTCA 2094
Db 1949 GAGAATTCAGATCGAAAGAGAGATCAGAATGGGTGATCATATCATCAGCAGACAGTCA 2008
QY 2095 AGCGCTTCTGCTACGTAGTCTACAAACAGAAAGGATTCAGGCAATTAACCTGCGCATCGGT 2154
Db 2009 AGCGCTTCTGCTCCGTAGCTCGAGAGAGAGATTCAGGCAATTAACCTGCTGTCAGCTGT 2068
QY 2155 GGAACATGGGTTTATACAAACTCTTCTTAAAGTAAACCTCGGAAGTCAATGACACAGACA 2214
Db 2069 GGAACACGGATTCATGCAAACTCTTCTTAAAGTAAACCTCGGAAGTCAATGACACAGACA 2128
QY 2215 TTTGAGAGACTTCTTATAAGATGATGATGAGATGGCTCTAGACCAAGAAATGTC 2274
Db 2129 TTTGAGAGAACTTCTTATAAGATGATGATGAGATGGCTCTAGATAAAGAAATGTC 2188
QY 2275 CAATAGCATGACACCTAGCCAGAGTCTGGTACAGAGACTTTCATGACAGTCAATCAACA 2334
Db 2189 GAGCAGCATGACGCCAGCCAGAGTCTGGTACGAGACTTTCATGACAGTCAATCAACA 2248
QY 2335 CCCCACCTTCAACACAGATGAGTCTGTGAACAAAGTTTGGAAAAGGACCGGAAACA 2394
Db 2249 CCCCACCTTCAACACAGATGAGTCTGTGAACAAAGTTTGGAAAAGGACCGGAAACA 2308
QY 2395 AGCTGGGCAAGCCAGGACATACCCAGGGAACAGTAAACAAATGGAAGCACTTCAAGA 2454
Db 2309 AGCCGACAAAGCCGGGCACTCTCAGGGAGAGCAACAAAGTGAAGCAGATCAAGA 2368
QY 2455 AAATAAGAAAGTGAAGAACAGGAGACCCACCAATTTGAGAGGCGACCCAGGAGTGTCTG 2514
Db 2369 GAGCAGAAAGTGAAGAACAGGAGACCCACCAATTTGAGAGGCGACCCAGGAGTGTCTG 2428
QY 2515 AGCTGATTAACCTTGAACAACTTCAACAAAGTGAAGTCTGCTAGACAAATACCTGAAA 2574
Db 2429 AGCTGGCCACCTCCCAAACTTCAACAAAGTGAAGTCTGCTAGATAAATACTGAAA 2488
QY 2575 ACAAATGCAATATACAACTTTTTCATGCGCATATGCGATGTTTACATGCGTGG 2634
Db 2489 ---AAATGCAATATACAACTTTCATGCGCATATGCGATGTTTACATGCGTGG 2543
QY 2635 AAATTCAGCTGAGTTCACCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 2694
Db 2544 AGTTCAACCGGTTCCCAAT-----TTAAGTCCGTGAGGAACCTTTCCAGCAGGCT 2597
QY 2695 TTTTTCCTA 2704
```

```
Db 2598 TTCTTCAGTA 2607
RESULT 10
RNSIIICNI
LOCUS R.norvegicus mRNA for semaphorin III/collapsin-1. linear ROD 29-JAN-1996
ACCESSION X95286
VERSION 1
KEYWORDS collapse-1; semaphorin III.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Giger, R.I.
AUTHORS Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2319)
AUTHORS Giger, R.I.
TITLE Direct Submission
JOURNAL
SUBMITTED (22-JAN-1996) R.I. Giger, Netherlands Inst. for Brain
Research, Meibergdreef 33, 1105 AZ, Amsterdam, NETHERLANDS
LOCATION/Qualifiers
FEATURES
1..2319
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="brain"
/dev_stage="adult"
1..2319
/codon_start=1
/product="semaphorin III/collapsin-1"
/protein_id="CAA64607.1"
/db_xref="GI:1171619"
/db_xref="SPTREMBL:O63548"
/translation="MGWFTGIACLFHWGILLTARANYANGKNNVRLKLSYKEMLESNN
VITNGLANSSYHTFLDEERSLYGAKDHFSPNLVNIKDFOKIWPVPSYTRDE
CKWAKILKECANPIKVLKAYNQTHLYACGTGAPHCITCTYIEVGHHPEDNIFKLQDS
HFENGKSPYDPKLLTASLIDGELYSGTADEFMRDFAIFRLTGLHHPHRTQDHS
RWLNDPRFISAHLLIPESDNPDQKVFYFFRENAIDGESHGKATHARICQIKNDPFGH
RSLVNWTTFLKARLICVSPFGNIDTDFDELQDVFLLMNSKDPKNPIVYGVFTSSNI
KGSACVMSMSDVRRVPLAGYAHRDGPNQWVPYQGRVPRPCTCPSKFGGPDST
KDLDDVTTPARSHPMYVPFINNRPIMIKTDVNTQFTQIIVDRVDAEDQYDVMF
IGTDVTVLKVSVPKETWHLLEEMTVFREPTTISAMELSTKQQLYIGSTAG
VAQPLHRCDIYGKACAECLLARDPYCAWDGSSCSRYPTTAKRTRRQIRNGDPLTH
CSDIQHDHNNHSHLEERIIYGVNSSTFLFECSPKQALVYQFORNEDEIRV
GDHIRTQGLLLSLQKDSNYLCHAVHGFMTLLKVTLEVIDTDEHLELLHOD
DGDGSKTKEMSSMTSPQKVMYRDFMOLINHPNLMTDFCFCEQVWKRDRKRRQRPQH
SQGSNNKWKMQESKGNRRTHFERAPRSV"
BASE COUNT 704 a 491 c 548 g 576 t
ORIGIN
Query Match 70.5%; Score 1909.4; DB 10; Length 2319;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 2070; Conservative 0; Mismatches 246; Indels 3; Gaps 1;
QY 200 ATGGCTGTTAACTAGGATTTCTGCTTTTTCGGGGAGTATTACTTACAGCAAGACA 259
Db 1 ATGGCTGTTTCACTGGGATTCCTCTTTTCTGGGGTATATTACTTACAGCAAGACA 60
QY 260 AACTATCAGATCGGAAGAACAAATGTCCAGGCTGAATTTATCTCAAGAAATGTTG 319
Db 61 AACTATCAAAACGGAAAGAACAAATGTCCAGACTGAATTTATCTGTAAGAAATGTTG 120
QY 320 GAATCCAAATATGTCATCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTTC 379
Db 121 GAATCCAAATATGTCATCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTTC 180
QY 380 CTTTGGATGAGGAACCGGAGTAGCTGTATCTTGGAGCAAGAGATCATATTTTCATTC 439
```


181	CTTCTGGATGAGGAACGGAGTAGACTGTATGTTGGAGCAAGGATCATATATTTTTCATTC	240
440	GACCTGGTTAATATCAAGAGATTTTCAAAAGATTTGTGGCCAGTATCTTATCACCCAGAAGA	499
241	AACTTGGTGNATATTAAAGATTTTCAAAAGATTTGTGGCCAGTATCTTATACCGAGAGA	300
500	GATGAATGCAAGTGGCGTGGAAAAAGACATCTCTGAAGAAATGTGTCTAAATTTTCATCAAGGTA	559
301	GATGAATGCAAGTGGCGTGGAAAAAGATATCTCTGAAGAAATGTGTCAAAATTTTCATCAAGGTC	360
560	CTTAAGGCATATATATCAGATCCTCTGTACGCCCTGTGMAAGGGGGCTTTTTCATCCAAAT	619
361	CTCAAGGCTTATAATCAGACTCATCTGTACGCCCTGTGGAACGGGGGCTTTTCCATCCAAAT	420
620	TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGGAGAACTCA	679
421	TGCACCTATATTGAAGTTTGGNCATCATCTCTGAGGAGCAACATTTTAACTTGCAGGACTCA	480
680	CATTTTGAAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGTCTGACAGCATCCCTT	739
481	CATTTTGAAAAACGGCTGTGGGAAGAGTCTTTATGACCCCAAACTTCTGTACTGCTCTCTT	540
740	TTAATAGATGGAGAAATTATACTCTGGAACTGCAGCTGATTTTATATGGGGCCAGACTTTGCT	799
541	CTAATAGATGGTGAAGTTATATCTCTGGAATCTCTGCGAATTTTCATGGGACAGAGCTTTGCT	600
800	ATCTTTCGGAACCTTTTGGGCCACCAACCCCAATCAGGACAGACAGCAGCATGATTCACAGTGG	859
601	ATCTTTCAGAACATTTGGGGCCACCATCTCCATCAGGACCGGACGACGACTCCCGTGG	660
860	CTCAATGATCCAAAGTTCAATTTAGTGGCCCACTCATCTCAGAGAGTGACAATCTCTGAAGAT	919
661	CTCAATGATCCTAGATTCATCAGTGGCCCATCTCATCTCAGAGAGTGATAACCCCTGAAGAT	720
920	GACAAAGTATACTTTTTCTCCGTGAAAAATGCAATAGATGAGGAAACACTCTGGAAGAAGCT	979
721	GACAAAGTATATTTTTTTCTTCCGAGAAAAATGCAATAGATGAGGAAACACTCTGGAAGAAGCC	780
980	ACTCACGCTAGAAATPAGTTCAGATATCAAGAAATGACATTTGGAGGGCACAGAGTCTGGTG	1039
781	ACTCATGCTAGATPAGTTCAGATATGCAAGATGACATTTGGAGGACACAGAGTCTTGTG	840
1040	AATAAATGGACAAACATTTCTCTCAAAAGCTCGTCTGATTTTGTCTCAGTGGCCAGGTCCAAATGGC	1099
841	AATAAATGGACAACTTTTCTCTGAAAGCACGCTGATTTGCTCTGTGTCAGGTCCCAATGGC	900
1100	ATTGACACTCATTTTGTATGAATGCAATGAGATGATTCCTTAATGAACTTTAAAGATCTCTAAA	1159
901	ATTGACACACATTTTGTATGAATTTACAGGATGATTTCTTAATGAACTTTAAAGATCTCTAAA	960
1160	AATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG	1219
961	ATYCCAATTTGTATGAGTGTTCACAAACATCAAGCAACATCTTTAAGGCTCTGTGTG	1020
1220	TGTATGTATAGCATGAGTGTGTGAGAAGGTTGTCTTGGTCCATATGCCCCACAGGAT	1279
1021	TGTATGTACAGTATGAGTGTATGAAGAGGTTGTCTTGGTCCATATGCTCACAGAT	1080
1280	GGACCCAACTATCAATGGGTGCCCTTATCAAGGAAGAGTCCCTCTATCCACGGCCAGGAAGCT	1339
1081	GGTCCCAACTATCAGTGGGTACCTTTACCAAGGAAGAGTCCCTTATCCACGACGAGAACT	1140
1340	TGTCACGACAAAACATTTGGTTGGTTTGTACTCTCAAGAGGACCTTCTCTGATGATGTATA	1399
1141	TGTCACGAGTAAACATTTGGTGGATTTGACTTCCAAAGGACCTTCTCTGATGATGTATA	1200
1400	ACCTTTTGCAAGAAGTCAATCCAGCCATGTACAATCTCAGTGTCTCTATGAAACAATTCGCCCA	1459
1201	ACCTTTTGCAAGAAGCCATCCAGCCATGTACAAACCCAGTGTCTCTATTAATATTCGCCCA	1260
1460	ATAGTGATCAAAACGGATGTAAATTTATCAAAATTTTACAAATTTGTGTGAGACCGAGTGGAT	1519
1261	ATAATGATCAAAACAGATGTAATTTATTCAGTTTACACAAATTTGTGTGAGACCGGTTGGAT	1320

QY	1520	GCAGAGATGGA	CAGTATGATGATGTTATTCGGAA	CAGATGTTGGGACCGTTCTTAAA	1579
DB	1321	GCAGAGATGGC	CAGTATGATGTTATGTTTCATCGAA	CGGATGTTGGAACCTGTCCTTAAA	1380
QY	1580	GTAGTTTCAAT	TCTCTAAGGAGACTTGATATGATTTAGAA	GAGGTTCTGCTGGAGAAATG	1639
DB	1381	GTGGTTTCAGT	CCCCAAGGAGACTTGGCATGACTTAGA	GAGGTTCTTCTGGAGAAATG	1440
QY	1640	ACAGTTTTCGG	AAACCGACTGCTATTTCACGAATGGAGCTT	TCCACTAAGCAGCAACAA	1699
DB	1441	ACAGTCTTCGG	AAACCAACAACTATTTCAGCAATGGAGCT	TCTTACTATAACAGCAACAA	1500
QY	1700	CTATATATTGG	TTCACCGGTGGGGTTGCCAGCTCCCTTTAC	CACCGGTGTGATATTTAC	1759
DB	1501	CTGTACATTTGG	CTCAACTGCTGGAGTAGCACAGCTTCTCTAC	ACCGGTTGTGACATTTAT	1560
QY	1760	GGGAAACGCTG	TGTGAGTGTCCCTCGCCGAGAGACCTTACT	TGTCTTTGGGATGTTCT	1819
DB	1561	GGCAAGCCTGT	GGGAATGCTGCTTGTCTGGGACCTTATTGT	GCTTGGGATGGGTCA	1620
QY	1820	GCATGTTTCTC	GTATTTTCCACTGCAAAAGAGACGCA	CAAGACGACAAGATATAAGAAAT	1879
DB	1621	TCATGCTACG	CTATTTTCTTACTGCAAAAGAGCGCACAA	GACGACAGGATATAAGGAAAT	1680
QY	1980	GGAGACCACTG	ACTACTGTTTCAAGACTT---ACACCATGAT	TAATACCATGGCCACAGC	1936
DB	1681	GGAGACCACTG	ACTACTGCTCCGACTTCGACATCATGAT	TAATCATCATGGGCACAGT	1740
QY	1937	CCTGAAGAGAG	AATCATCTATGTGTAGAGAAATAGTAG	CACATTTTGGAAATGCAGTCGG	1996
DB	1741	CTTGAAGAGAG	AATCATCTATGGAGTAGAAATAGTAG	CACATTTCTTGGAAATGCAGTCGG	1800
QY	1997	AAGTCGAGAG	AGCGTGTGTTATTTGGCAATTTCCAGAG	GCGAAATGGAAGCGAAAGAA	2056
DB	1801	AAGTCACAGAG	AGCGTTTGGTATATTGGCAATTTTCCAG	GCGAAATGGAAGCGAAAGAA	1860
QY	2057	GAGATCAGAGT	GCATCATATCATCAGCAGACAGATCA	AGGGCTTCTGCTACGTAGTCTTA	2116
DB	1861	GAGATCAGAGT	GGTGATCATATCATCAGCAGACAGAC	GGGGTCTCTGCTTCGTAGTCTTG	1920
QY	2117	CAACAGAAGGA	TTCAGGCAATTAACCTCTGCCATGCGGT	GGAAACATGGGTTTCATACAACT	2176
DB	1921	CAGAAAGGAAT	TCAGGCAATTAACCTGTGTCAAGCGGT	GAGCAGGTTTTCATGCAAACT	1980
QY	2177	CTTCTTAAGGT	AACCTTGGAGTTCATGTACACAGAGCAT	TTTGGAGAACTTCTTCTATAAA	2236
DB	1981	CTTCTTAAGGT	AGTCAACCTTGGAGTTCATGTATACAGAA	CATTTGGAGGAACTTCTTCTATAAA	2040
QY	2237	GATGATGATGG	AGTGTCTTAGACCAAGAAATGTCCTCAAT	ATAGCATGACCTAGCCAG	2296
DB	2041	GATGATGATGG	AGTGTCTTAGACCAAGAAATGTCCTCAAT	ATAGCATGACCTAGCCAG	2100
QY	2297	AAGTCTGTGTA	CAGAGACTTCATGTCAGCTCATCAAC	CCCCCACTTCAACAGTGGAT	2356
DB	2101	AAGTCTGTGTA	CAGAGACTTCATGTCAGCTCATCAAC	CCCCCACTTCAACAGTGGAT	2160
QY	2357	GAGTTCGTGTA	CAAGTTTGGAAAAGGGACCGGAAACAA	AGCTTCGGCAAGGGCCAGGCAT	2416
DB	2161	GAGTTCGTGTA	CAAGTTTGGAAAAGGGACCGGAAACAA	AGCTTCGGCAAGGGCCAGGCAT	2220
QY	2417	ACCCAGGGAA	CAGTAAACAAATGGAAGCACTTACAGA	AAATTAAGAAAGGTAGAACAGG	2476
DB	2221	TCTCAAGGGAG	CAGCAACAAGTGAAGCACTGCAGAG	AGCAAGNAAGGTAGAACAGG	2280
QY	2477	AGGACCCAGAA	TTTGGAGGGCACCCAGAGTGTCTGA		2515
DB	2281	AGGACCCAGAA	TTTGGAGGGCACCCAGAGTGTCTGA		2319

RESULT 11
MUSCOSE
LOCUS

MUSCOSE 2319 bp mRNA linear ROD 13-FEB-2002

DEFINITION	Mus musculus collapsin/semaphorin III mRNA, complete cds.	
ACCESSION	L41541	
VERSION	L41541.1	GI:18654369
KEYWORDS	collapsin/semaphorinIII.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 2319)	
TITLE	Kimura,T. and Fishman,M.C.	
JOURNAL	cDNA sequence of mouse Collapsin/SemaphorinIII	
FEATURES	Unpublished	
source	Location/Qualifiers	
	1..2319	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/db_xref="taxon:10090"	
	/dev_stage="adult"	
	1..2319	
	/note="putative"	
	/codon_start=1	
	/product="collapsin/semaphorin III"	
	/protein_id="AA177611.1"	
	/db_xref="GI:18654370"	
	/translation="MGWFTGIACLFWGVLLTARANYANGKNNVRLKLSYKEMLESNVITFNGLANSSYHTFLDDEERSLYVGAKDHFNFNLVNIKDFQKIVPVSYTRRDECKWAGKDLKECANFTKYLEAYNQTHLYACGTGAGHPICITYIEVGHPEDNIFKLQDHSHPENGKSPYDKLLTASLLDGLYSCTAADFMGRDPAIFRTLGHHPHPIRTEOHDRLWNPRIISAHLIPESONPEDDKVYFFRENAIDGESHGKATHARIQIICKNDFGHRSLVNWTFLKARLICSPVPGNGIDHDELQDVLNNSDKPNPIVIGVFTTSNIRFGSAVMTFSRSDRRVFLGVAHNRDGNPOVYGRVYPRPTGCTSKPTGGFDVDFKOLPDVITFARSHFAMTNPVPIINRIMIKTDVNYQTVIIVDRVDAEDQYDVMFIGTDVTVLKVYSPKETHWLDLEVLLEMTVFREPTTISAMELSTKQQLYIGSTAGVAQLPLHRDIDYKGAECCLARDPYCAWDSSSKRYPPTAKRRTRRDIRNGDPLTHCSDLOHNDHGHPSLEERIIYGVENSTFLECSKPSORALVYVQFQRENDEKKEIRMGDHIIRTEQGLLRSLQKDSGNLYLCHAVEHGMOTLLKLVILEVIDTEHLEHLHKDDGGSKIKEMSSMTSPSKVTRDFMQLNHPNLNMTDEFCEQVWKRKORRORRPHSGSSNNKWKHMESKKGKRRNRTHFERAPRSV"	
BASE COUNT	703 a 513 c 549 g 554 t	
ORIGIN		
Query Match	69.5%; Score 1883.8; DB 10; Length 2319;	
Best Local Similarity	88.6%; Pred. No. 0;	
Matches 2054; Conservative	0; Mismatch 262; Indels 3; Gaps 1;	
Qy	200	ATGGGCTGGTTAACTAGGATTGTCTCTTTCTGGGGAGTATTACTTACAGCAAGACGA 259
Db	1	ATGGGCTGGTTCACTGGGATGCTCTTTCTGGGGTGTATTACTTACGCCAGACGA 60
Qy	260	AACTATCAGAATGGGAAGAACAAATGTGCCAAGCTGAAATTATCTTACAAAGAAATGTTG 319
Db	61	AACTATGCAACGGAAAGAACAAATGTGCCAAGCTGAAATTATCTTACAAAGAAATGTTG 120
Qy	320	GAATCCAAAGTGTGATCACTTTCAATGGCTTGGCCAAAGCTCAGATTATCATCCTTC 379
Db	121	GAATCCAAAGTGTGATCACTTTTAAATGGCTTGGCCAAAGCTCAGATTATCATCCTTC 180
Qy	380	CTTTTGGATGAGCAAGGATAGGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTC 439
Db	181	CTTTTGGATGAGCAAGGATAGGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTC 240
Qy	440	GACCTGGTTTAAATCAAGATTCTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGA 499
Db	241	AACTTGGTGAACATTAAAGATTTCAAGATTGTGTGGCCAGTATCTTACACCAAGAGA 300
Qy	500	GATGAATGAAGTGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTTCAATCAAGTTA 559
Db	301	GATGAATGAAGTGGCTGGAAAGATATCTCGAAAGATGTGCTAAATTTCAATCAAGTTC 360
Qy	560	CTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCAATT 619
Db	361	CTGAGGCTTATAATCAGACTCACTTGTATGCTGTGGAACTGGGGCTTTTCATCAATT 420
Qy	620	TGCACCTCATTTGAAATTTGGACATCATCTCGAGGACAAATATTTTTTAAGCTGAGAACTCA 679
Db	421	TGCACCTTATATTGAAGTTGGACATCATCTCGAGGACAAATATTTTTTAAGCTGAGAACTCA 480
Qy	680	CATTTTGAACGGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACACATCCCTT 739
Db	481	CATTTTGAACGGCGCTGGGAAGAGCCCTTATGATCCAAACTACTGACTGCTCTCTT 540
Qy	740	TTAATAGATGGAGAAATTAATCTCTGGAACCTGACGCTGATTTTTATGGGGGAGACTTTGCT 799
Db	541	CTAATAGACGGTGTGCTGCTGGAACCTGCTGGGACTTCTATGGGACGGGACTTCGCT 600
Qy	800	ATCTTCCGAACCTCTTGGGACCAACCAATCAGACAGACAGACAGATGATTCAGGTGG 859
Db	601	ATCTTCCGAACCTCTTGGGACCAATCAGGACCAATCAGGACGAGCAGTACTCCGGTGG 660
Qy	860	CTCAATGATCAAAAGTTCATTAGTCCACCTCATCTCAGAGAGTACAATCTCTCAAGAT 919
Db	661	CTCAATGATCCTAGATTCATTAGTCCACCTCATCTCAGAGAGTACAATCTCTCAAGAT 720
Qy	920	GACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGGCT 979
Db	721	GACAAAGTATATTTTCTTCCGAGAAATGCAATAGACGAGAGAACACTCTGGAAGGCT 780
Qy	980	ACTACGCTAGATAGGTGAGATGCAAGAAATGACTTTTGGAGGACACAGAACTGTGGTG 1039
Db	781	ACTCATGCTAGATAGGTGAGATGCAAGAAATGACTTTTGGTGGACACAGAAAGTCTTGTG 840
Qy	1040	AATAAATGGAACAATCTCTCAAGCTCGTCTGATTCTGCTCAGTCCGAGGTCCAAATGGC 1099
Db	841	AATAAATGGAACAATCTCTTAAAGCACGCTGATTCTCTGTGCCGGTCCCAATGGC 900
Qy	1100	ATTGACACTCATTTTGTGAACTGAGGATGTATTCTTAATGAACTTTTAAAGATCTCTAAA 1159
Db	901	ATTGACACCACTTTTGTGAACTGAGGATGTATTCTTAATGAACTCTTAAAGATCTTAAA 960
Qy	1160	AATCAGTTGTATATGGAGTGTATTACGACTTCCAGTAAACATTTTCAAGGATCAGCCGTG 1219
Db	961	AATCCGATGCTATGGAGTGTTCACAACATCAAGCAACATCTTTTAAAGGATCTGCTGTG 1020
Qy	1220	TGATGTATAGCATGATGTGAGAGGTTGTTCTTGTGCTCATATGCCCACAGGAT 1279
Db	1021	TGATGTATAGCATGATGTGAGAGGTTGTTCTTGTGCTCATATGCTCACAGAGAT 1080
Qy	1280	GGACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGCGGCGAGAACT 1339
Db	1081	GGTCCCACTATCAGTGGTGCCTTACCAAGAGAGTCCCTTATCCAGCGGCGAGAACT 1140
Qy	1340	TGTCCCAAGAAAACATTTGGTGTGCTCTACAAAGGACCTTCTCTGATGATGTTATA 1399
Db	1141	TGTCCCAAGAAAACATTTGGGCGGATTTGACTCCCAAGAGGACCTTCTCTGATGTCATA 1200
Qy	1400	ACCTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTTATGAACATCGCCA 1459
Db	1201	ACTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTTATGAATATCGCCCG 1260
Qy	1460	ATAGTGTCAAAAACGAGTGTAAATTTATCAATTTTACAAAATTTGCTAGACCGAGTGGAT 1519
Db	1261	ATCATGATCAAAAACAGATGTAAATTTATCAGTTTCACAAAATTTGTTAGACCGAGTGGAT 1320
Qy	1520	GCAGAAGATGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGTCTCTTAAA 1579
Db	1321	GCAGAAGATGCCAGTATGATGTTTATGTTTATCGGAACAGATGTTGGAAACCGTCTCTTAAA 1380
Qy	1580	GTAGTTTCAATTCCTAAGGACACTTGGTATGATTTAGAAAGAGTGTCTCTGGAAGAAATG 1639
Db	1381	GTGGTTTCAGTCCCAAGGAGACTTGGCATGACCTAGAAAGATTTCTTCTGGAAGAAATG 1440
Qy	1640	ACAGTTTTTTCGGGAACCGACTGCTATTTCACAAATGGAGCTTTCACCTAAGCAGCAAA 1699
Db	1441	ACCGTCTTCCGGGAACCAACAACTATTTCGGCAATGGAGCTTTTACTTAAACAGCAACAG 1500
Qy	1700	CTATATATTGTTCAACGGCTGGGGTGTGCCAGCTCCCTTTTACACCGGTGTGATATTTAC 1759

Db	1501	CTGTACATTTGGCTCAACTCGGGAGTGGCACAGCTTCCTCTACACCGCTGTGACATCAT	1560
Qy	1760	GGGAAACGCTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTCTGGGATGGTTC	1819
Db	1561	GGCNAAGCCTGTGCAGAAATGCTGCTCGCTCGGAGCCCTTACTGTGCTCGGATGGGTCC	1620
Qy	1820	GCATGTCTTCGCTATTTTTCCCACTGCAAAAGAGACGCAAGAGACAAAGATATAAGAAAT	1879
Db	1621	TCATGCTCACGCTATTTTTCTTACTGCAAGAGGCGCAAGACAGCAAGATATAAGGAAT	1680
Qy	1880	GGAGACCCTGACTCACTGTTTCAGACTT---ACACCATGATATCAACATGGCCACAGC	1936
Db	1681	GGAGACCCTGACTCACTGCTCTGACTTTGCAGCACCATGATATCATCATGGGCCAGC	1740
Qy	1937	CCTGAACAGAGAAATCATCTATGGTGTAGAAATAGTAGACATTTTTTGGAAATCGACTCCG	1996
Db	1741	CTTGAAGAGAAATCATCTATGGAGTGGAAACAGTAGTAGTACATCTTTGGAAATGCAGTCCG	1800
Qy	1997	AAGTTCGAGAGAGCGCTGGTCTATTGGCAATTCAGAGGCGCAATGAAGAGCGAAAGAA	2056
Db	1801	AAGTCAAGAGAGCCTTGGTATATTGGCAATTTTCAGAGAGAAATGAAGATCGAAAGAG	1860
Qy	2057	GAGATCAGAGTGGATGATCATATCATCAGCACAGATCAAGGCCCTTCTGTACGTAGTCTA	2116
Db	1861	GAGATCAGAATGGTGTATCATCATCAGGACAGAAACAAGGGCTCTGTCTCCGTAGCCTG	1920
Qy	2117	CAACAGAAGAAATTCAGGCAATTACCTCTGCGATCGGTGAAACATGGGTTTCATACAACT	2176
Db	1921	CAGAAGAAGAAATTCAGGCAATTAACCTGTGTCTGCTGTGGAACACGGAAATTCATGCAAACT	1980
Qy	2177	CTTCTTAAAGTTAACCTCGGAAGTCATTGACACAGAGCATTTTGAAGAACTTCTTTCATAA	2236
Db	1981	CTTCTTAAAGTTAACCTCGGAAGTCATTGACACAGAACATTTTGAAGAACTTCTTTCATAA	2040
Qy	2237	GATGATGATGGAGATGCTCTTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAG	2296
Db	2041	GATGACGATGGAGATGCTCTAAGATAAAAGAAATGTTCGAGCAGCATGACGCCACGCCAG	2100
Qy	2297	AAGTCTGTGTACAGAGACTTCATGCACTCATCAACACCCCAATCTCAACACGATGGAT	2356
Db	2101	AAAGTCTGTGTACCGAGACTTCATGTGACCTCATTTAAACCACCCCACTGAAACGATGGAT	2160
Qy	2357	GAGTTCGTGAAACAAAGTTTCGAAAGGACCGGAAACAAACGTGGCAAAAGGCCAGAGACAT	2416
Db	2161	GAGTTCGTGTGAAACAAAGTTGTGAAAGGACCGAAAGCAACGCCGACAAGAGCCGGGGCAC	2220
Qy	2417	ACCCAGGGGAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAACAGG	2476
Db	2221	TCTCAAGGGAGCAGCAACAGTGAAGCAATCAAGAGAGCAAGAAAGGTAGAAACAGG	2280
Qy	2477	AGGACCCACGAAATTTGAGAGGGCACCCAGAGGTGTCGA	2515
Db	2281	AGGACCCACGAGTTTGAGCGGGCACCCAGAAAGTGTCTGA	2319

RESULT	12
GGU02528	
LOCUS	3263 bp mRNA linear VRT 01-NOV-1993
DEFINITION	Gallus gallus collapsin mRNA, complete cds.
ACCESSION	U02528
VERSION	U02528.1 GI:410078
KEYWORDS	.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
	Phasianinae; Gallus.
REFERENCE	1 (bases 1 to 3263) Luo, X., Raible,D. and Raper,J.A. Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones <i>J Cell Biol</i> 75 (2), 217-227 (1993)
AUTHORS	
TITLE	
JOURNAL	

Db 577 TTCATCCCATGTCATTAATGAGTTGGAGCCATCTCAGGACAACTTTTAGGA 636
QY TGGAGAACTCACATTTTGAACCGCCGCTGGAGAGTCCATATGACCCCTAAGCTGCTGA 728
Db 637 TGGAGATTACATTTTGAACCGCCGAGGAGAAAGCCCTTACGATCCTAAATTTGCTGA 696
QY GAGCATCCCTTTTAATAGATGAGAAATTAATCTCTGGAACTCGAGCTGATTTTATGGGCG 788
Db 697 CAGCTTCTCTTTTAGTAGTGAGAAATGTACTCTGGCACAGCAGCAGACTTTCATGGGCA 756
QY GAGACTTTGCTATCTTCCGAACTCTTGGGCACACACCCCAATCAGGACAGAGCAGCATG 848
Db 757 GAGACTTTGCCATTTCCGAACTCTGGGCATCACCATCCAATCAGAACAGAGCAGCATG 816
QY ATTCCAGGTGGCTCAATGATCAAAAGTTCAATAGTGCCCACTCATCTCAGAGAGTGACA 908
Db 817 ACTCCCGGTGGCTAAACGATCCAGATTTATTTAGTGCTCACCTGATACCTGAAAGTGATA 876
QY ATCTTGAAGATGACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACTACT 968
Db 877 ATCCGAGAGATGACAAATCTATTTCTCTTCTGTAATGCGATTTGATGGAGAGCACA 936
QY CTGGAAGAGCTACTCAGCTAGAAATAGGTGAGATGCAAGAAATGACTTTGGAGGCGACA 1028
Db 937 CTGGAAGAGCTACTCATGCGAATAGGCGAGATCTGTAAGAAATGACTTTGGTGAGCACA 996
QY GAGGTCTGCTGTAATTAATGGAACAACATTTCTCAAAAGCTCGTCTGATTTGCTCAGTGCCAG 1088
Db 997 GGAGCTTGTGTAACAAATGGCAACCTTTCCTCAAGCAGCTGATTTGTTCTGTGGCAG 1056
QY GTCCAAATGGCAATGACACTCAATTTTGTATGAACTGAGAGATGATTTCTTAATGAATTTA 1148
Db 1057 GACCAATAGTATGACACACACTTTGATGAATTAAGAGCAGCTGTTCTTAATGAATCTTA 1116
QY AAGATCTTAAATATCCAGTTGATATGAGTGTGAGAGGCTTACAGTTTCCAGTAAATATCTTCAAGG 1208
Db 1117 AGGACCTTAAATATCCAGTTGATGAGTGTGAGAGGCTTACAGTTTCCAGTAAATATCTTCAAGG 1176
QY GATCAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
Db 1177 GATCAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
QY CCCACAGGATGAGCCAACTATCAATGAGTCCCTTATCAAGAAAGAGTCCCTATCCAC 1328
Db 1237 CCCACAGGATGAGCCAACTATCAATGAGTCCCTTATCAAGAAAGAGTCCCTATCCAC 1296
QY GGCAGGAGTGTCCAGCAAAACATTTGGTGGTTTGGTCTTCAAGAGACCTTCTG 1388
Db 1297 GGCAGGAGTGTCCAGCAAAACATTTGGTGGTTTGGTCTTCAAGAGACCTTCTG 1356
QY ATGATGTTTAACTTTGCAAGAGTCAATCCAGCCATGTACAACTCCAGTGTTCCTATGA 1448
Db 1357 ATGAGTTTAACTTTGCAAGAGTCAATCCAGCCATGTATACCCAGTATTCCTCCATCA 1416
QY ACAATGCCCCAATGATGATCAAAACGAGTGAATTAATCAATTTTACAAATTTGCTGAG 1508
Db 1417 ACAGCGGTCCCAATCATGATCAAAACGAGTGTAGATTAACAGTTTCAACAGATAGTAG 1476
QY ACCGAGTGGATGCAAGATGCAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1568
Db 1477 ATCGAGTATGATGCAAGATGCTCAATACGATGATGATGATGATGATGATGATGATGATG 1536
QY CGGTCTTTAAAGTGTGTTCAATTTCTAAGGAGACTTGGTATGATTTTGAAGAGTTCG 1628
Db 1537 CAGTTCTTAAGTGTGTTCCATTCCTAAGGAGACTTGGCAATGATTAAGAGAGTCTTGC 1596
QY TGGAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGGCTTTCCACTA 1688
Db 1597 TGGAGAAATGACAGTCTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGATTTCCACA 1656
QY AGCAGCAACACTATATATGTTTCAAGGCTGCGGCTGCGGCTCCCTTTACACCGGT 1748

Db 1657 AACGCAACAGCTCTACATTTGGCTCAGCCACGGAGTTTACAGCTTCTCTTACACCGCT 1716
QY GTGATATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCCCTTACTGTGCTT 1808
Db 1717 GTGATGTCTATGGAAGAACATGTGCTGAGTGTGCTTTCGAAGAGACCCCTTACTGCGCT 1776
QY GGGATGTTCTGCAATGTTCTGCTATTTTCCACTGCAAGAGAGCGCACAGACACAAG 1868
Db 1777 GGGATGTTCTATCTTGTCTCAGCTTACTTTCCACTGCTAAGAGAGCTACTAGGCGACAAG 1836
QY ATATAAGAAATGGAGACCCACTGCTCAGCTGTTGAGCTTACACCATGATATCACC--- 1925
Db 1837 ACATCCGAATGGAGATCCCTCCTCAGCTGCTCAGCTGCTGAGCATCATGATACCCAA 1896
QY ATGGCCACAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTACACATTTTGG 1985
Db 1897 GTGGTCAGACCTCGAGGAGAAAGATTAATTTATGGATAGAGAAATAGCAGCACTTTTCTTG 1956
QY AATGAGTCCGAAGTCCGAGAGAGCGTGTGCTATTTGSCAATTCAGAGGCGGAATGAG 2045
Db 1957 AGTGAGTCCAAAATCTCAACGTGCTATCTGTCGCAATTCAGAGCAAAATGATG 2016
QY AGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGSCCTTCTGC 2105
Db 2017 ACCAAGAGTTGAGATAAAGTGGATGATCGAATGATCCGACAGAACAGGSCCTTATGC 2076
QY TACGTAGTCTACAACAGAGGATTCAGGCAATTTACTCTGCCATCGCGTGGAACTATGGT 2165
Db 2077 TACGAAGTCTTCAACGGAGAGACTCTGCTATTTATTTTGTGCTGCTGAGCATGCT 2136
QY TCATCAAACTCTTCTTAAGTAACTTGAAGTCACTTGAAGTCACTGACGACATTTGGAAGAC 2225
Db 2137 TCATTCAACTTGTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2196
QY TTCTTCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2285
Db 2197 TGCTCCATGAAGAGAGACCCGATGCTCCAGACCAAGATGCTACCAATAGCATGA 2256
QY CACTAGCAGAGAGTCTGTGACAGAGACTTTCAGCAGCTCATCAACACACCCCAATCTCA 2345
Db 2257 CTCTAGTCAAAAGATCTGTTATAGAGACTTTCAGCAGTAAATCAATCACCACCAATCTCA 2316
QY ACAGATGATGATGTTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGCTCGGCAAA 2405
Db 2317 ACACAAATGATGAATTTCTGTGAGCAGGTTTGGAAAGAGACCGCAAGCAGCGCGCAAA 2376
QY GGCAGGACATACCCAGGGAACATCAATGGAAGCCTTACAGCAAAATAGAAAG 2465
Db 2377 GGCCTGCCAATGCGCAGGTGAACCAAAATAAGTGGAGCACTTACAGAGAAATAAAAG 2436
QY GTAGAAACAGGAGACCCACCAATTTGAGAGGACCCAGAGTGTCTGAGCTGATTCAC 2525
Db 2437 GTAGAAATAGAGAGACCCATGAATTTGAGAGGACCCAGAGTGTCTGAGCTACATTCAC 2496
QY CTCTAGAAACCTCAAAACAGTGAAGACTTGTGCTGAGCAATTAATCTGGAAGAAACAAATGCAA 2585
Db 2497 CTCTAGAAACCTCATCTAAGTAGAGAACTTGTATAGACAGATAACTTGGAAAAAATAGCAA 2556
QY TATACATGAATTTTTCATGCGCAATTTGAGTGTGATGATGATGATGATGATGATGATGATG 2645
Db 2557 TATACATGAAC-TTTTTCATGCGCAATTAAGTGAATGTTTACAGAGTGTGAAAGCTCAACA 2615
QY AGTTTCCACCAATTAATAATCCATGAGTAACTTTCTTAATAGGCTTTTTC 2701
Db 2616 ATTTCCACCGGTTTAAATACATATATTTCTAATCTTCTTAGTAGTCTTCTGTC 2671

RESULT 13

MUSSEMAIII

LOCUS

DEFINITION

ACCESSION

VERSION

MUSSEMAIII 1998 bp mRNA linear ROD 03-AUG-1995

Mus musculus semaphorin III (Sema III) mRNA, partial cds.

L40484

L40484.1 GI:703189

Db 781 GCAACACATGCGAATTCGACAGCTATGCAAGATGATTTGGTGGCCACAGAGCTTA 840
Qy 1037 GTGATAAATGAGACAACTTCCTCAAGCTCGTCTGATTTGCTAGTGCAGGTCAAT 1096
Db 841 GTTAACAAAATGGACTACTTTCTTAAAGCTCGCTAAATTTGTTAGTGCCTGGTCTTAAT 900
Qy 1097 GCAATGACACTCATTTTCATGAACCTGACAGATGATTCCTTAATGAACCTTAAAGATCCT 1156
Db 901 GGCATTTGATACACATTTTATGATGAACCTACAGATGTTGTTTGTGATGAACCTCAAAGATCCA 960
Qy 1157 AAAAAATCCAGTTGTATATGAGAGTTTACGACTTCAGTAAACATTTTCAAGGGATCAGCC 1216
Db 961 AAAAAATCCAGTTGTATATGAGAGTTTACGACTTCAGTAAACATTTTCAAGGGATCAGCA 1020
Qy 1217 GTGTGATGATAGATGATGATGAGAGGTTTCTTGGTCCATATGCCACAGG 1276
Db 1021 GTCTGTATGATGATGATGATGAGAGGTTTCTTGGTCCATATGCCACAGG 1080
Qy 1277 GATGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCCAGGA 1336
Db 1081 GATGGTCCAAATTTATCAGTGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCCAGGA 1140
Qy 1337 ACTGTGCCAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACCTTCTCGATGATGT 1396
Db 1141 ACTTGGCCCTAGGCAGACTTTTGGTGGCTTTGACTCAAAAGGATCTCCCTGATGAAGTT 1200
Qy 1397 ATAACTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACATCGC 1456
Db 1201 ATTATGTTTGCAGAAAGTCAACCGGCCATGTCAATCTCTGTTTTCCTCAATCAATCGT 1260
Qy 1457 CCAATAGTGCATCAAAACGAGTGTAAATTTACAAATTTACAAATTTGCTAGACCGAGTG 1516
Db 1261 CCATCATTAATAAAACGAGTGGATTAATTTTACAAATTTAGTGGTGGATCGAGTG 1320
Qy 1517 GATCGAAGAGTGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTT 1576
Db 1321 GAAGCTGAGGATGGTCAATACGATGTCATGTTTATTTGGAACCTGATATGGGGACCATTTTG 1380
Qy 1577 AAAGTATGTTCAATCTTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGNA 1636
Db 1381 AAAGTATGTTCTGTTTCCCAAGGAGACTCGGACTGATTTTGAAGAAGTTCTTACTGGAAGAA 1440
Qy 1637 ATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACATAAGCAGCAA 1696
Db 1441 ATGACTGTATTTAGGAGGCAACTGCTATTTTACGCAATGGAGATTTCAAAAGCAACAA 1500
Qy 1697 CAACATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTT 1756
Db 1501 CAACATATATTTGGCTCGTCAGTGGGTGTTTCCCAACTCCCGCTTCCAGCGTGTGATTT 1560
Qy 1757 TAGCGGAAGAGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGATGGT 1816
Db 1561 TATGGAAGAGCTTGTGCTGAGTGTTCCTTGGCCTTGGCAGAGACCTTATTTGCTTGGATGGA 1620
Qy 1817 TCTGATGTTCTGCTGCTATTTTCCACTGCAAGAGAGCGACAGAGCAAGATATAAGA 1876
Db 1621 TCTTATGCTACGCTTACTTCTGCTTAAAGAGCTACTAGGGCCAGATATCAGG 1680
Qy 1877 AATGGAGACCCACTGACTTCACTGTTTCCAGACTTACACCAT- - -GATAATCACCATGGCCAC 1933
Db 1681 AAGGGGATCCCTTAACATATTTTCCAGACCTTACAGCATCAAGATGATCCACATAGACAA 1740
Qy 1934 AGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGAATGTCAGT 1993
Db 1741 AGCCTAGAGAAAGATAATCTATGTTGTGGAGAACAGTAGACATTTCTCTGGAATGCAGC 1800
Qy 1994 CGGAAGTCCAGAGAGCGGTGTTCTATTTGCAATTCAGAGCGCAATCAAGAGCGAA 2053
Db 1801 CCTAATACAGCGTGCCTAGTTTCTGCGAGTTCCAGAGCAAAATGAAGAAAAAG 1860
Qy 2054 GAAGAGATCAGAGTGGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAGT 2113

Db 1861 GATGAGATAAAGTGGATGAACGCAAAATAAAGACAGAACATGGCCTTCTTCTGCGCAG 1920
Qy 2114 CTACAACAGAGGATTTCAGGCAATTTACCTCTGCCATGCGGTGGACATGGGTTTCATACAA 2173
Db 1921 CTAAAGAAAGAGAGATTTCAGGCATATATTATTGCAATGCAAGTATGAGTATGTTTTCATGAG 1980
Qy 2174 ACTCTTCTTAAGGTAAACCTCGGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCAT 2233
Db 1981 ACCCTTCTCAAGTAACTCTCGAGATTAATGATACAGAACATTTAGATGAGCTTCTTCAC 2040
Qy 2234 AAAGATGATGATGAGATGGCTCT- - -AAGACCAGAAAGAAATGCTCAATAGCATGACCT 2290
Db 2041 AAAGATGATGAGGAGGAGACAGTACAAACACAAAGAACCTTCAACAGATATGCTTCCA 2100
Qy 2291 AGCCAGAGGTCTGTTACAGAGACTTTCATGAGCTCATCAACCCCACTCAACAG 2350
Db 2101 ACTCAAAAGATGATGATAGAGCTTTTATGAGTTTAAACCCCAAAATCTTAAACACA 2160
Qy 2351 ATGATGATGTTCTGTGAACAAAGTTTGGAAAGGACCGAAACAAACGTCGCGCAAGGCCA 2410
Db 2161 ATGATGATGTTTGTGAGCAGGTGTGAAAGGACCGCAAGCAACGTCGCGCAAAATAAT 2220
Qy 2411 GGACATATCCCAAGGAAACAGTAAACAAATGGAAGCACTTTACAGAAATAAGAAAGGTAGA 2470
Db 2221 GGAATGTCCAGTTTAGCAACACAAATGGAAGCACCTTACAGGAAACCAAGAGGGGTAGA 2280
Qy 2471 AACAGGAGGCCACAGCAATTTGAGAGGGCACCCAGGAGTGTCTGA 2515
Db 2281 AATAGGAGGCCCATGAATTTGAGAGGGCACCAAGGAGTGTCTGA 2325

RESULT 15
111900
LOCUS 111900 1481 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 1 from Patent US 5416197.
ACCESSION I11900
VERSION I11900.1 GI:909343
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1481)
AUTHORS Raper,J.A. and Luo,Y.
TITLE Antibodies which bind human collapsin
JOURNAL Patent: US 5416197-A 1 16-MAY-1995;
FEATURES Location/Qualifiers
source
1..1481
/organism="unknown"
BASE COUNT 454 a 299 c 345 g 383 t
ORIGIN

Query Match 52.2%; Score 1415.2; DB 6; Length 1481;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 899 GAGAGTGACAACTCCTGAAGATGACAAAGTACTATTTTCTTCGTTGAAAATGCAATAGAT 958
Db 53 GAACATGACAACTCCTGAAGATGACAAAGTACTATTTTCTTCGTTGAAAATGCAATAGAT 112
Qy 959 GGAGAACACTCTGGAAAAGCTTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACATTT 1018
Db 113 GGAGAACACTCTGGAAAAGCTTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACATTT 172
Qy 1019 GGAGGGCACAGAGTCTGGTCAATAAATGGAACAACTTCTCAAGCTCGTCTGATTTGC 1078
Db 173 GGAGGGCACAGAGTCTGGTCAATAAATGGAACAACTTCTCAAGCTCGTCTGATTTGC 232
Qy 1079 TCAGTGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAATGACAGAGTATTTCTCTA 1138
Db 233 TCAGTGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAATGACAGAGTATTTCTCTA 292
Qy 1139 ATGAATTTAAAGATCTTAAATAATCCAGTCTGTATATGAGTGTTTACAGCTTCCAGTAAC 1198

Db 293 ATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATATGGAGTGTATTACGACTTCCAGTAAC 352
QY 1199 ATTTTCAAGGGATCAGCCGTGRTATGTATAGCATGAGTGTGAGAGGGTGTTCCTT 1258
Db 353 ATTTTCAAGGGATCAGCCGTGRTATGTATAGCATGAGTGTGAGAGGGTGTTCCTT 412
QY 1259 GGTTCATATGCCACAGGATGGACCCAACTATCAATGGTGCCCTTATCAAGGAAGTTC 1318
Db 413 GGTTCATATGCCACAGGATGGACCCAACTATCAATGGTGCCCTTATCAAGGAAGTTC 472
QY 1319 CCTATCCACGGCCAGGACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAAAG 1378
Db 473 CCTATCCACGGCCAGGACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAAAG 532
QY 1379 GACCTTCTGTATGATGTTTATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTG 1438
Db 533 GACCTTCTGTATGATGTTTATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTG 592
QY 1439 TTTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAAATTTCAATTTACAAA 1498
Db 593 TTTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAAATTTCAATTTACAAA 652
QY 1499 ATTGTCGTAGACCGAGTGGATCAGAAGATGACAGATGATGATGTTTATGTTTATCGGAACA 1558
Db 653 ATCGTCGTAGACCGAGTGGATCAGAAGATGACAGATGATGATGTTTATGTTTATCGGAACA 712
QY 1559 GATGTGGGACCGTTCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAA 1618
Db 713 GATGTGGGACCGTTCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAA 772
QY 1619 GAGGTTCTGCTGGAAGAAATGACAGTTTTTGGGAAACCGACTGTCTATTTTACGCAATGGAG 1678
Db 773 GAGGTTCTGCTGGAAGAAATGACAGTTTTTGGGAAACCGACTGTCTATTTTACGCAATGGAG 832
QY 1679 CTTTCCACTAAGCAGCAACAATATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCT 1738
Db 833 CTTTCCACTAAGCAGCAACAATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCT 892
QY 1739 TTACACCGGTGTGATATTTACGGGAAAGCGTGTCTGAGTGTGCTCGCCCGAGACCCCT 1798
Db 893 TTACACCGGTGTGATATTTACGGGAAAGCGTGTCTGAGTGTGCTCGCCCGAGACCCCT 952
QY 1799 TACTGTCTTGGGATGGTTCTGCATGTTCTGCTATTTTCCACTGCAAAAGAGACGCACA 1858
Db 953 TACTGTCTTGGGATGGTTCTGCATGTTCTGCTATTTTCCACTGCAAAAGAGACGCACA 1012
QY 1859 AGACGACAAGATATAGAAATGAGACCCACTGCTCACTGTTTCACTTACAGCTTACACCATGAT 1918
Db 1013 AGACGACAAGATATAGAAATGAGACCCACTGCTCACTGTTTCACTTACAGCTTACACCATGAT 1072
QY 1919 AATCACCATTGGCCACAGCCCTGAAGAGAGAATCATCTATGTTGTAGAGATAGTAGCACA 1978
Db 1073 AATCACCATTGGCCACAGCCCTGAAGAGAGAATCATCTATGTTGTAGAGATAGTAGCACA 1132
QY 1979 TTTTGTGAATGAGTCCGAAGTCCGAGAGCGCTGTTCTATTTGGCAATTCAGAGCGGA 2038
Db 1133 TTTTGTGAATGAGTCCGAAGTCCGAGAGCGCTGTTCTATTTGGCAATTCAGAGCGGA 1192
QY 2039 AATGAAGCGGAAAGAGAGATCAGAGTGGATGATCATCATCAGCAGAGATCAAGGC 2098
Db 1193 AATGAAGCGGAAAGAGAGATCAGAGTGGATGATCATCATCAGCAGAGATCAAGGC 1252
QY 2099 CTTCCTGCTACGTAGTCTCAACAGAGAGATTCAGGCAATTTACCTCTGCGATGCGGTGAA 2158
Db 1253 CTTCCTGCTACGTAGTCTCAACAGAGAGATTCAGGCAATTTACCTCTGCGATGCGGTGAA 1312
QY 2159 CATGGGTTTATCAAAACTCTTCTTAAGGTAAACCTTGGAAAGTCATTGACACAGAGCATTTG 2218
Db 1313 CATGGGTTTATCAAAACTCTTCTTAAGGTAAACCTTGGAAAGTCATTGACACAGAGCATTTG 1372
QY 2219 GAAGAACTTCTTATGAAGATGATGAGATGAGTCTTAAGACCAAAAGAAATGTCCAAT 2278
Db 1373 GAAGAACTTCTTATGAAGATGATGAGATGAGTCTTAAGACCAAAAGAAATGTCCAAT 1432

QY 2279 AGCATGACACCTAGCCAGAGGTTCTGGTACAGAGACTTTCATGCAGCTC 2326
Db 1433 AGCATGACACCTAGCCAGAGGTTCTGGTACAGAGACTTTCATGCAGCTC 1480

Search completed: July 31, 2003, 15:53:21
Job time : 9588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:47:35 ; Search time 5029 Seconds
(without alignments)
13092.230 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttcatgatg.....agggttttttctctaacc 2709

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	26.7	723	12	BI870437 603395690
2	715.2	26.4	3317	11	AK052671 Mus muscu
3	715.2	26.4	3672	11	AK028900 Mus muscu
4	712	26.3	4248	11	AK031704 Mus muscu

5	674.8	24.9	679	12	BM711125
6	663.8	24.5	702	12	BM723807
7	620.2	22.9	2960	11	AK043634
8	620.2	22.9	3154	11	AK033597
9	620.2	22.9	3292	11	AK037034
10	617	22.8	3287	11	AK048364
11	594.6	21.9	729	14	CD349263
12	590.4	21.8	900	14	CB196368
13	588.6	21.7	889	10	BF667677
14	581	21.4	775	13	BU703431
15	567.4	20.9	699	14	CA327319
16	543.8	20.1	604	12	BI183160
17	537.8	19.9	773	13	BUS15916
18	524.2	19.4	3436	11	AK014333
19	522.4	19.3	3884	11	AK034239
20	510	18.8	570	10	BE033028
21	506.8	18.7	510	14	CD216784
22	504.2	18.6	3394	11	AK053632
23	501.8	18.5	698	13	BU461875
24	485.8	17.9	884	10	BF700780
25	475.4	17.5	559	14	CA872840
26	463.4	17.1	807	14	CB595903
27	458.6	16.9	2596	11	AK053115
28	433.2	16.7	660	10	BB612039
29	439.4	16.2	497	14	CA395261
30	428.2	15.8	731	13	BU708579
31	397.6	14.7	917	13	BQ884924
32	388.4	14.3	747	12	BQ018643
33	385.6	14.2	1809	11	AK051165
34	382	14.1	382	13	EX283116
35	378.4	14.0	493	14	CB725110
36	366.4	13.5	914	14	CA488964
37	364.6	13.5	787	14	CB518391
38	360.6	13.3	511	13	BU445100
39	353	13.0	441	14	CA871124
40	351.4	13.0	600	12	BI989239
41	344.4	12.7	721	2	HS084323
42	341.4	12.6	692	13	BQ447821
43	336	12.4	839	10	BF667143
44	331.2	12.2	773	28	BH265538
45	325.2	12.0	394	14	CB776561

ALIGNMENTS

RESULT 1
BI870437
LOCUS 603395690F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405519 5',
DEFINITION 723 bp mRNA linear EST 11-OCT-2001
mRNA sequence.
ACCESSION BI870437
VERSION BI870437.1 GI:16044110
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 723)
NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbe-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM2034 row: c column: 24
High quality sequence stop: 723.

FEATURES
source

Location/Qualifiers

1. .723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405519"
/issue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

BASE COUNT 225 a 148 c 159 g 191 t

ORIGIN

Query Match 26.7%; Score 723; DB 12; Length 723;

Best Local Similarity 100.0%; Pred. NO. 8.7e-173;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 CAAGGATTTTCAAGATGTGTGCCAGTATCTTACACCGAAGAGATGATGCAAGTG 513
DB 1 CAAGGATTTTCAAGATGTGTGCCAGTATCTTACACCGAAGAGATGATGCAAGTG 60
QY 514 GCGTGGAAAAGACATCCTGAAAAGATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAA 573
DB 61 GCGTGGAAAAGACATCCTGAAAAGATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAA 120
QY 574 TCAGACTCATTTGTAGCGCTGTGGAAACGGGGCTTTTCATCAATTTGCACCTACATTGA 633
DB 121 TCAGACTCATTTGTAGCGCTGTGGAAACGGGGCTTTTCATCAATTTGCACCTACATTGA 180
QY 634 AATTCGACATCATCTCTGAGCAATATTTTAAAGTCTGAGAACTCACATTTTGAACCG 693
DB 181 AATTCGACATCATCTCTGAGCAATATTTTAAAGTCTGAGAACTCACATTTTGAACCG 240
QY 694 CGGTGGGAAGAGTCCATATGACCTTAAGCTGTGCAGCATCCCTTTTAATAGATGGAGA 753
DB 241 CGGTGGGAAGAGTCCATATGACCTTAAGCTGTGCAGCATCCCTTTTAATAGATGGAGA 300
QY 754 ATTATCTCTGAACTGACGCTGATTTTATGGGGCGAGACTTTGCTATCTTCCGAACCTCT 813
DB 301 ATTATCTCTGAACTGACGCTGATTTTATGGGGCGAGACTTTGCTATCTTCCGAACCTCT 360
QY 814 TGGGACACCAACCAATCAGACAGCAGCATGATTCAGGTGGCTCAATGATGCCAA 873
DB 361 TGGGACACCAACCAATCAGACAGCAGCATGATTCAGGTGGCTCAATGATGCCAA 420
QY 874 GTTCATTAGTGGCCACCTCATCTCAGAGAGTGACATCTGGAAGATGACAAAGTATACTT 933
DB 421 GTTCATTAGTGGCCACCTCATCTCAGAGAGTGACATCTGGAAGATGACAAAGTATACTT 480
QY 934 TTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTCTGAAAGACTACTACGCTAGAT 993
DB 481 TTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTCTGAAAGACTACTACGCTAGAT 540
QY 994 AGGTGAGATGTCAGATGACTTTGGAGGGGACAGAGATCTGTGTGATTAATGACAAAC 1053
DB 541 AGGTGAGATGTCAGATGACTTTGGAGGGGACAGAGATCTGTGTGATTAATGACAAAC 600
QY 1054 ATTCTCTAAAGCTCGCTGATTTGCTCAGTGCAGGTCCAAATGCAATGACACTCATTT 1113
DB 601 ATTCTCTAAAGCTCGCTGATTTGCTCAGTGCAGGTCCAAATGCAATGACACTCATTT 660
QY 1114 TGATGAATGCGAGATGATTTCTTAATGAACTTTAAAGATCCTAAAAATCCAGTTGTATA 1173
DB 661 TGATGAATGCGAGATGATTTCTTAATGAACTTTAAAGATCCTAAAAATCCAGTTGTATA 720
QY 1174 TGG 1176
DB 721 TGG 723

RESULT 2

AK052671

LOCUS

DEFINITION

AK052671 3317 bp mRNA linear HTC 05-DEC-2002
Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D630018P03 product:SMAPHORIN 3D PRECURSOR homolog
[Homo sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitho, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

1..3317

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6J"

/db_xref="FANTOM DB:D630018P03"

/db_xref="taxon:10090"

/clone="D630018P03"

/tissue_type="kidney"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

548..2881

/note="unlabeled protein product; SEMAPHORIN 3D PRECURSOR homolog [Homo sapiens] (SWISSPROT|O95025, evidence: FASTA, 92.8%ID, 100%length, match=2331)

Putative"

/codon_start=1

/protein_id="BAC35092.1"

/db_xref="GI:26342871"

/translation="MNVTKDSEPSRQDLHLFHAMWMLIMVLFLPVTESKNIPLKLYKDLLLNTCPFLGSEGLDFQTLLEDEGILLIGAKDHVFLLSLVLDKNF KKIYPAKAKERVELCKLAGDANAEANFIRVLPYKNTKHYVCGTGAFLPLCYIDL GANKEELIFKLDTHLSGRCLKPFPQPPASVMTDEHLYSGTASDFLGKDTAFTRS LGLMQDHSIRTDISEHHLNGAKFIFGTPIDTNPDDDKIYFFPFLSSQEGSTDR SILSRVGRVCKNDVGGQSLNKWTFLKARLCSIPGSDGATHFDELQDYLPLPR DERNPVYGVPTTSSIFPKGSACVYSMDIRAVFNGPYAHKESADHRWQYDGRIPY PRPGTCSKTYDPLIKSTRDPDVISFIRHPVNYKSVIPVAGATFKRINDVIRLT QIVVDHVAEDQGDVGLFQDITGVLVKVSISKEKNMEVEVLEELQVFKHPTAILN MELSKQOVLVSGWGLQVLSHRCDTYGRACADCLARDPYCAMDNGNCSRYPAIS KRRARQDVQDPTQCDWIEDISTSHETADEKVI FGIENSTFELCI PKSQOASVW YIORSQDHRLEKPDRI IKTQYGLIRLSRQKDSGWYCKAOEHTFHTIVKLTLN VIENQMENTORAOYQGOVKDLAERSLRYKDYIQLSSNFSFLDQYCEOMWYKEKR RQNKGSFKWHEMKKRNRRHRLDELQRSVAT"

CDS

939 a 756 c 807 g 815 t

Query Match 26.4%; Score 715.2; DB 11; Length 3317;

Best Local Similarity 62.1%; Pred. No. 1.6e-170;

Matches 1220; Conservative 0; Mismatches 723; Indels 21; Gaps 5;

BASE COUNT

ORIGIN

275 AAGACAATGCGCAAGCTGAATATCTACAAAGAAATGTTGGAATCCAAACAATGTG 334

Db 662 AAACAAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAACACCTGT 721

Qy 335 ATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCCTTTTCGATGAGAA 394

Db 722 ATCCCCCTTTTGGGTTTCATCAGAGAGATGGAATTCAGACTCTTCTTTTGGATGAGGAG 781

Qy 395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATAATTTTCATTCGACCTGGTTAA---T 451

Db 782 AGGGGCATCTGCTCCTAGGAGCCAAAGACCATGTCTCTCTGCTCAGTCTGTTGACTTGTG 841

Qy 452 ATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTTACACAGAAGAGATGAATGCAAG 511

Db 842 AACAGAAGATTTTAAAGAAGATTTATTTGGCTCTGCTCAAAAGAACGAGTGGAGCTATGAAA 901

Qy 512 TGGCTCGGAAAGACATCTCTGAAGAATGTCTAATTTTCATCAAGGTACTTTAAGGCATAT 571

Db 902 TTAGCTGGGAAGATGCAATGACAGATGTGCAAAATTTTCATCCTGTGCTTCAACCCCTAT 961

Qy 572 AATCAGACTCACTTGTAGCGCTGTGGAAACGGGGGCTTTTTCATCAAAATTTGACCTACAT 631

Db 962 AATAAGACTCAGCTTTAGCTGTGTGGAACTGGAGGGTTTCATCCGCTGTGTGGGTACAT 1021

Qy 632 GAAATTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAANA 691

Db 1022 GATCTCGGCGCCAAAGAGGAGAACTCATATTTAAACTAGACAGCAACACTGGAGTCT 1081

Qy 692 GGCGCTGGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751

Db 1082 GGCAGACTGAAATGTCCCTTTGATCCTCAGCAGCTTTTGTCTCAGTAATGACAGATGAG 1141

Qy 752 GAAATTATCTCTGAACTGAGCTGATTTTATGGGCGAGACTTTGTCTATCTCCGAACT 811

Db 1142 CACCTCTACTCTGAAACAGCTTCTGATTTCTTGGCAAGACACTGCAATTCACAGGTCT 1201

Qy 812 CTTGGGC-----ACCACCAACCCAAATCAGGACAGCAGCATGATTCAGGTGGCTC 862

Db 1202 CTGGGGCTTAATGACGAGCACCACCATTCATCAGAACTGACATTTTCAGAGCACCCTGGCTC 1261

Qy 863 AATGATCCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGCATCTCGAAGATGAC 922

Db 1262 AATGAGCAAAATTTATCGGAACATTCCTCCATTCAGCACTTATTAATCCAGATGATGAT 1321

Qy 923 AAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGAGAAACACTCTCGAAAAAGCTACT 982

Db 1322 AAATATATTTCTTCTTTCGAGAAATCATCCAGAGAGCAGTACTTCTTGACAGAGCAT 1381

Qy 983 CACCTAGAATAGTCCAGATATGCAAGAATGACTTTTGGAGGGCACAGAAGTCTCGTGAAT 1042

Db 1382 CTTTCAAGAGTGGGAAGAGTTTGTGAAGATGATAGTAGTGGGCAACGAAGTCTGATAAAC 1441

Qy 1043 AATGGACACATCTCCTCAAGCTCGTCTGATTTGTCTCAGTCCAGCTCCAAATGGCAT 1102

Db 1442 AAATGGCAACTTTTCTTAAAGGCAAGACTGATTTGTCTCGATTCCTGGAAGCGATGGGGCA 1501

Qy 1103 GACACTCATTTTGTATGAATGACAGGATGATTTCTTAATGAATTTTAAAGATCTTAAAGAT 1162

Db 1502 GATACCACTTTGATGNACTCCAGATATTTTACTTACTCCTAGGAGATGAAGAAT 1561

Qy 1163 CCAGTTGTATAGGAGTGTTCAGACTTCAGTAACATTTTCAAGGGATCAGCGGTGT 1222

Db 1562 CCTGTAGTATATGAGTCTTTTACCAACACAGCTCCATCTTCAAAGGCTCTGCTGTCTGT 1621

Qy 1223 ATGTATAGCATGATGATGAGAGAGGTGTTCTTGGTCCATATGCCCCACAGGATGGA 1282

Db 1622 GTGTACAGCATGGCTGATATCCGAGCAGTCTTTTAATGGTCCCTTATGCTCATAGGAAGT 1681

Qy 1283 CCCAACTATCAATGGGTGGCTTATCAGGAAGAGTCCCTTATCCAGCCCGCAGGAACCTGT 1342

Db 1682 GCTGACCATCGTGGGTGCAATATGATGAAGGATACCTTACCCCGCAGCTTGAACCTGT 1741

Qy 1343 CCAGCAAAAACATTG---GTGGTTTGTACTCTACAAAGGACCTTCTCTGATGATGTTATA 1399

Db 1742 CCAAGCAAAACCTATGACCACCATGATTAAGTCCACCCGAGCTTTCAGACGATGTTATC 1801

Qy 1400 ACCTTTCAGAGAAGTATCAGGCCATGTACAATCCAGTGTGTTCTTATGAACAATFCGCCCA 1459

Db 1802 AGTTTCATAAGGGGGCACCTGTGTATAGTCCGTGATAGTCCAGTGGGAGGAGCAGCCG 1861

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

location/Qualifiers
1..3672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:4732468K20"
/db_xref="taxon:10090"
/clone="4732468K20"
/tissue type="skin"
/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
329..3662
/note="unnamed protein product; SEMAPHORIN 3D PRECURSOR
homolog [Homo sapiens] (SWISSPROT|095025, evidence: FASTV,
92.8%ID, 100%length, match=2331)
putative"
/codon_start=1
/protein_id="BAC26185.1"
/db_xref="GI:26324862"
/translation="MNVTKDNPSSQDLHPHAWMLIMTVLFLPVTETSKONIPR
KKIYPAKREVELCKLAGDANACANFIVLOPKYKNTVYVCGTGAFFHPCGYIDL
GANKEBLLFKLTDHNSGRLLKCPDPQPPASVWTDHLYSGTASDFLKDRAFTSR
LGLWQDHSIKTDISEHNLNGAKFIGTPIPDYNNPDDDKIYFFRESSQBSGTSOR
SILSRGVKNDGQSRSLINKWTFILKARLICSIPSGDGDTHFDELQDIYLLPTR
DERNPVYGVFTTSSIFKSGAVCVSMADIRAVFNGPYAHKESADHRWQYDGRIPY
PRPGTSPSKTDPILKIPDPDDIVSFIRHPVMYKSVYVAGAPFPRINVDYRLT
QIVDVHVARDQYDVMFLGTDIGTVLKVSIKSKNMWEVLEELQVFKHPITALIN
MELSILKQQLVGVSDGLVQLSLRCDYTGKACDCLARDPYCAMDGNACSRVAPIS
XRRARRQVKTGPDITQCDIEDISIHETADEKVIIFGFNFSTFLECIPKSGQASVW
VIQSGDEHRELPKDERIKTDYGLLRSLQKDSGMVYCKAQEHTFIHTIVKLIN
VIQSGDEHRELPKDERIKTDYGLLRSLQKDSGMVYCKAQEHTFIHTIVKLIN
RQRNGSGPKWKMOMKRRNRHRLDELQSVAT"

CDS

/note="unnamed protein product; SEMAPHORIN 3D PRECURSOR
homolog [Homo sapiens] (SWISSPROT|095025, evidence: FASTV,
92.8%ID, 100%length, match=2331)
putative"
/codon_start=1
/protein_id="BAC26185.1"
/db_xref="GI:26324862"
/translation="MNVTKDNPSSQDLHPHAWMLIMTVLFLPVTETSKONIPR
KKIYPAKREVELCKLAGDANACANFIVLOPKYKNTVYVCGTGAFFHPCGYIDL
GANKEBLLFKLTDHNSGRLLKCPDPQPPASVWTDHLYSGTASDFLKDRAFTSR
LGLWQDHSIKTDISEHNLNGAKFIGTPIPDYNNPDDDKIYFFRESSQBSGTSOR
SILSRGVKNDGQSRSLINKWTFILKARLICSIPSGDGDTHFDELQDIYLLPTR
DERNPVYGVFTTSSIFKSGAVCVSMADIRAVFNGPYAHKESADHRWQYDGRIPY
PRPGTSPSKTDPILKIPDPDDIVSFIRHPVMYKSVYVAGAPFPRINVDYRLT
QIVDVHVARDQYDVMFLGTDIGTVLKVSIKSKNMWEVLEELQVFKHPITALIN
MELSILKQQLVGVSDGLVQLSLRCDYTGKACDCLARDPYCAMDGNACSRVAPIS
XRRARRQVKTGPDITQCDIEDISIHETADEKVIIFGFNFSTFLECIPKSGQASVW
VIQSGDEHRELPKDERIKTDYGLLRSLQKDSGMVYCKAQEHTFIHTIVKLIN
VIQSGDEHRELPKDERIKTDYGLLRSLQKDSGMVYCKAQEHTFIHTIVKLIN
RQRNGSGPKWKMOMKRRNRHRLDELQSVAT"

BASE COUNT

ORIGIN 1091 a 787 c 855 g 939 t

Query Match 26.4%; Score 715.2; DB 11; Length 3672;
Best Local Similarity 62.1%; Pred. No. 1.7e-170;
Matches 1220; Conservative 0; Mismatches 723; Indels 21; Gaps 5;
275 AAGACATGCGCAAGCTGAATATTCCTACAAAGAAATGTTGGAATCCAAATGTG 334
443 AAACAAATATTCAAAGACTCAAGCTAACCTACAAAGACTTGTCTTCAACACCTGT 502
335 ATCACTTTCAATGGCTTGGCAAGCTCCAGTTCATATTCATCTTCTTGGATGAGAA 394
503 ATCCCTTTTGGGTTCATCAGAGGATGGAATTCAGACTCTTCTTTGGATGAGAG 562
395 CGGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAA -T 451
563 AGGGGCACTGCTCTTAGGAGCAAGAACCACTGTCTTCTGCTCAGTCTGTTGACTTG 622
452 ATCAAGGATTTTCAAGATTTGTGGCGCATATCTTACACAGAGAGATGAATGCAAG 511
623 AACAGGATTTTAAAGAGATTTATTTGGCTCTCTGCAAAAGAACGAGTGGAGCTATGAAA 682
512 TGGGCTGGAAGACATCCTGAAGATGCTATATTTTCATCAAGGATCTTAAGGCATAT 571
683 TTAGCTGGAAGATGCCAATGCAGATGTGCAATTTTCATCCGTGTGCTTCAACCCCTAT 742
572 AATCAGACTCACTGTGTAGCGCTGTGGACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
743 AATAAGACTCAGTTTACGTGTGTGGAATCGAGCGTTTCATCCGCTGTGTGGGTACATT 802
632 GAAATTGACATCATCTCTGAGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAC 691

803 GATCTCGGCGCCAAACAGGAGAACTCATATTTAAACTAGACACGCAACACTGGAGTCT 862
692 GGGCGTGGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGA 751
863 GGCAGACTGAAATGTCCTTTGATCCTCAGCAGCCTTTTGTCTCAGTAAATGACAGATGAG 922
752 GAATTTACTCTGAACTGACGCTGATTTTATGTTGGGCGAGACTTTTGTCTATCTTCGAACT 811
923 CACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGACACTGCTGATTTCAAGGCTCT 982
812 CTTGGGC-----ACCACCCCAATCAGGACAGCAGCAGCATGATTTCCAGGTGGCTC 862
983 CTGGGCTTAATGACGAGCACCATTTCCATCAGAACTGACATTTGACAGCACCATCTGGCTC 1042
863 AATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCTCGAAGATGAC 922
1043 AATGAGCAAAATTTATCGGAACATTTCCCATTCAGACACCTATATATCCAGATGATGAT 1102
923 AAAGTATATTTTCTTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAAGAGCTACT 982
1103 AAAATATATTTCTTCTTTCGAGAAATCATCCAGGAAGGACGACTACTTCTGACAGAACAT 1162
983 CACGCTAGATAGCTCAGATATGCAAGAAATGACTTTTGGAGGCGACAGAGTCTGCTGAAT 1042
1163 CTTTCAAGAGTTGGAAGAGTTTGTGAAGAAATGATGAGTGGGCAACGAAAGTCTGATAAAC 1222
1043 AAATGGACAACATTCCTCAAGCTCGTCTGATTTTGTCTCAGTGCAGGTCCTCAATGGCAAT 1102
1223 AAATGGACAATTTCTTAAGGCAAGACTGATTTTGTCTCGATTTCTTGAAGCGATGGGCA 1282
1103 GACACTATTTTGAATGACGAGGATGATTTCTTAATGAACTTTTAAAGATCTTAAAGAT 1162
1283 GATACCATTTTGAATGAACTTCAAGATATTTACTTACTTCTCCCTAGCAGAGATGAAAGAA 1342
1163 CCAGTTCTATAGAGTGTTCAGACTTCCAGTAACTTTTCAAGGATCAGCGCTGTGT 1222
1343 CCTGTATATGAGTGTCTTTACCAACACGACTTCTTCAAGGCTCTGCTGTCTGT 1402
1223 ATGTATAGCATGAGTGTGAGAGGCTTCTTCTGCTCATATGCCCACAGAGGATGGA 1282
1403 GTGTACAGCATGGCTGATATCCGAGCAGTCTTTAATGGTCCCTATGCTCATAGGAAGT 1462
1283 CCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCGCAGGAATTTGT 1342
1463 GCTGACCATCGTGGGTGCAATATGATGGAAGGATACCTTACCCCGACCTTGAACGCTGT 1522
1343 CCAGCAAAACATTTG---GTGGTTTGTACTCTCAAGAGGCTTCTCTGATGATTTATA 1399
1523 CCAAGCAAAACCTATGACCCCACTGATTAAGTCCACCGAGACTTTCCAGAGCTGTTTATC 1582
1400 ACCTTTGCAGAGAGTCAATCCAGCCTATGTAATCCAGTGTTCCTATGAACAAATCGCCCA 1459
1583 AGTTTCATAGGCGGACCTCTGATGATTAAGTCCGTGTTACCCAGTGGCAGGAGCACC 1642
1460 ATAGTGTCAAAAAGGATGAAATTTCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
1643 ACCTTCAAGAGATCAACGCTGATTTACAGCTGACGAGATAGTGGTGGATCACTGCTGTC 1702
1520 GCAGAGATGACAGTATGATGTTTATGTTTATCGGAACAGATGTTGGACCGTCTTAAA 1579
1703 GCTGAAGCGGCGAGTATGATGTCATGTTTCTCGGAACAGACATTTGGAACAGTCTCTGAAA 1762
1580 GTAGTTTCAATTCCTAAGGAGACTTGTGATGATTTTGAAGAGGTTTCTGCTGGAAGAAATG 1639
1763 GTTGTGAGCATCTCAAGGAGAGAGTGG---AATATGGAAGAGGTCGTACTGGAGAGCTT 1819
1640 ACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTTCCACTAAGCAGCAACA 1699
1820 CAGGATTTCAAGCACCAACAGCTATCTTGAACATGAGTGTGCTGCTGAAGCAGCAACAG 1879
1700 CTATATATTCGTTCAACGGCTGGGTTGCCAGCTTCCCTTTACACCGGTGTGATTTTAC 1759
1880 TTGTACGTTGTTCTGGGATGGATTGTCAGCTCTCTTTCGACAGATGCGACACTTAC 1939

Db	2007	TCCCAACAAGCCTCTGTTGAGTGGTACATCCAGCGGTACAGAGATGAGCATCGAGAGGAG	2066
Qy	2060	ATCAGAGTGGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTACAA	2119
Db	2067	TTGNAACCTGATGNAAGGATCATCAAACTGACTATGGGCTACTGATTCGAAGTCTGCAG	2126
Qy	2120	CAGAAGGATTCAAGCAATTACCTCTGCCATGCGGTGGAAATGGGTTTCATACAAACTCTT	2179
Db	2127	AAGAAGGATTCGGATGTATTACTGCAAGACACAGGAGCACACTTTCATCCACACCATA	2186
Qy	2180	CTTAAGGTACCCCTGGAGTCAATTGACACAGAGCATTTTGAAGA	2223
Db	2187	GTGAAGCTGACTTTGAATGTCTATTGAGAATGAACAGATGAAAA	2230

RESULT 5					
BM711125					
LOCUS	BM711125	679 bp	mRNA	linear	EST 28-FEB-2002
DEFINITION	UI-E-DX1-agv-i-12-0-UI.r1 UI-E-DX1 Homo sapiens cDNA clone UI-E-DX1-agv-i-12-0-UI 5' mRNA sequence.				

ACCESSION	BM711125
VERSION	BM711125.1
KEYWORDS	GI:19024383
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 679)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)

```

Genetics (www.resgen.com)".
Seq primer: M13 Reverse.
Location/Qualifiers
1..679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX1-agv-i-12-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DX1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: Ecor I; Site_2: Not I;
UI-E-DX1 is a normalized cDNA library containing the
following tissue(s): fetal eyes. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA,
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGAAATCAGA. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye

```

BASE COUNT	242 a	133 c	161 g	142 t	1 others
ORIGIN					
Query Match	24.9%; Score 674.8; DB 12; Length 679;				
Best Local Similarity	99.6%; Pred. No. 1.6e-160;				
Matches 676; Conservative	0; Mismatches 3; Indels 0; Gaps 0;				
QY	1964	GAGAATAGTAGCACATTTTGGAAATCCAGTCCGAAAGTCGAGAGCGGCTGGTCTATTGG	2023		
DB	1	GAGAATAGTAGCACATTTTGGAAATCCAGTCCGAAAGTCGAGAGCGGCTGGTCTATTGG	60		
QY	2024	CAATCCAGAGCGAAATGAAGCGAAAGAGATCAGAGTGGATGATCATATCATC	2083		
DB	61	CAATCCAGAGCGAAATGAAGCGAAAGAGATCAGAGTGGATGATCATATCATC	120		
QY	2084	AGGACAGATCAAGGCTTCTGCTACGTAGTCTACACGAAGGATTCAGGCAATTACCTC	2143		
DB	121	AGGACAGATCAAGGCTTCTGCTACGTAGTCTACACGAAGGATTCAGGCAATTACCTC	180		
QY	2144	TGCGATCGCGTGGAACATCGGTTTCATACAAATCTTCTTAAAGTAAACCTTGAAGTCAAT	2203		
DB	181	TGCGATCGCGTGGAACATCGGTTTCATACAAATCTTCTTAAAGTAAACCTTGAAGTCAAT	240		
QY	2204	GACACAGACATTTGGAAGAACTTCTTCAAAAGATGATGATGGAGATGGCTCTAAGACC	2263		
DB	241	GACACAGACATTTGGAAGAACTTCTTCAAAAGATGATGATGGAGATGGCTCTAAGACC	300		
QY	2264	AAAGAAATGTCCATAGCATGACACCTAGCCAGAGGCTCTGGTACAGAGACTTCATGCAG	2323		
DB	301	AAAGAAATGTCCATAGCATGACACCTAGCCAGAGGCTCTGGTACAGAGACTTCATGCAG	360		
QY	2324	CTCATCAACACCCCAATCTCAACACGATGGATGAGTCTGTGAAACAAGTTTGGAAAAAGG	2383		
DB	361	CTCATCAACACCCCAATCTCAACACGATGGATGAGTCTGTGAAACAAGTTTGGAAAAAGG	420		
QY	2384	GACCGAAAAACACGTGGCAAGGCGGACGACATACCCAGGGAACAGTAAACAATGGAAG	2443		
DB	421	GACCGAAAAACACGTGGCAAGGCGGACGACATACCCAGGGAACAGTAAACAATGGAAG	480		
QY	2444	CACCTACAGAAATTAAGAAAGGTAGAAACAGAGGACCCACGAATTTGAGAGGGCACCC	2503		
DB	481	CACCTACAGAAATTAAGAAAGGTAGAAACAGAGGACCCACGAATTTGAGAGGGCACCC	540		
QY	2504	AGGAGTGTCTGAGCTGCATTACTCTAGAAAACTCAACAAGTAGAAAACTTGCTCTAGACA	2563		
DB	541	AGGAGTGTCTGAGCTGCATTACTCTAGAAAACTCAACAAGTAGAAAACTTGCTCTAGACA	600		
QY	2564	ATAACTGGAAAAACAAATGCAATATACATGAACCTTTTTTCATGGCATATGTGGATGTTT	2623		
DB	601	ATAACTGGAAAAACAAATGCAATATACATGAACCTTTTTTCATGGCATATGTGGATGTTT	660		
QY	2624	ACAAATGGTGGAAATTCAG	2642		
DB	661	ACAAATGGTGGAAATTCAG	679		
RESULT 6					
BM723807					
LOCUS	702 bp mRNA linear EST 01-MAR-2002				
DEFINITION	UI-E-E01-aiX-04-0-UI.r1 UI-E-E01 Homo sapiens cDNA clone				
ACCESSION	UI-E-E01-aiX-04-0-UI 5', mRNA sequence.				
VERSION	BM723807				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.				
TITLE	Bonaldo, M.F., Lennon, G. and Soares, M.B.				
	Normalization and subtraction: two approaches to facilitate gene discovery				

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
1. 702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E01-aix-o-04-0-UI"
/tissue_type="fetal eye"
/dev_stages="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."

BASE COUNT 248 a 145 c 164 g 145 t
ORIGIN

Query Match 24.5%; Score 663.8; DB 12; Length 702;
Best Local Similarity 99.3%; Pred. No. 1e-157;
Matches 698; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

Qy 1920 ATCCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTGTAGAGAAATAGTAGCACAT 1979
Db 1 ATCCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTGTAGAGAAATAGTAGCACAT 59

Qy 1980 TTTTGGATGAGTCCGAGTTCGAGAGAGCGTGGTCTATTGGCAATTCAGAGGGCGAA 2039
Db 60 TTTTGGATGAGTCCGAGTTCGAGAGAGCGTGGTCTATTGGCAATTCAGAGGGCGAA 119

Qy 2040 ATGAAGACGGAAGAGAGATCAGAGTGGATGATCATATCATCAGGACATCAAGGCC 2099
Db 120 ATGAAGACGGAAGAGAGATCAGAGTGGATGATCATATCATCAGGACATCAAGGCC 179

Qy 2100 TTCTGCTACGTAGTCTACAACAGAGGATTCAGGCAATTAACCTCTGCCATCGCGTGGAA 2159
Db 180 TTCTGCTACGTAGTCTACAACAGAGGATTCAGGCAATTAACCTCTGCCATCGCGTGGAA 239

Qy 2160 ATGGGTTTCATCAAACTCTTTAAGGTAACCCCTGGAGTCATTGACACAGAGCATTTGG 2219
Db 240 ATGGGTTTCATCAAACTCTTTAAGGTAACCCCTGGAGTCATTGACACAGAGCATTTGG 299

Qy 2220 AAGAACTCTTCATAAAGATCATGATGAGTGGTCTTAAGACCAAGAAATGTCGAATA 2279
Db 300 AAGAACTCTTCATAAAGATCATGATGAGTGGTCTTAAGACCAAGAAATGTCGAATA 359

Qy 2280 GCATGACACCTAGCCAGAGGTCTGGTACAGAGACTTTCATGACAGTCTATCAACACCCCA 2339
Db 360 GCATGACACCTAGCCAGAGGTCTGGTACAGAGACTTTCATGACAGTCTATCAACACCCCA 419

Qy 2340 ATCTCAACACGATGGATGAGTCTGTGAACAGTTTGGAAAAGGACCGAAAACAGCTC 2399
Db 420 ATCTCAACACGATGGATGAGTCTGTGAACAGTTTGGAAAAGGACCGAAAACAGCTC 479

Qy 2400 GGCAAGGCCAGGACATACCCAGGGAACAGTAACTGAAGCACTTCAAGAAAATA 2459
Db 480 GGCAAGGCCAGGACATACCCAGGGAACAGTAACTGAAGCACTTCAAGAAAATA 539

Qy 2460 AGAAAGGTAGAAAACAGGAGGCCACCGAA-TTTGAGAGGGCCACCCAGAGTGTCTGAGCT 2518
Db 540 AGAAAGGTAGAAAACAGGAGGCCACCGAA-TTTGAGAGGGCCACCCAGAGTGTCTGAGCT 599

Qy 2519 GCATTACTCTAGAAAACCTCAACAGTAGAAAACCTTGCCTAGACAATTAACGGAAAACA 2578
Db 600 GCATTACTCTAGAAAACCTCAACAGTAGAAAACCTTGCCTAGACAATTAACGGAAAACA 659

Qy 2579 AATCAATATACATGAAC-TTTTTTCATGGCATTATGTGGATG 2620
Db 660 AATCAATATACATGAAC-TTTTTTCATGGCATTATGTGGATG 702

RESULT 7
AK043634
LOCUS
DEFINITION
Mus musculus 10 days neonate cortex cDNA, RIKEN full-length
enriched library, clone:AB30014F15 product:SEMAPHORIN 3B PRECURSOR
(SEMAPHORIN A) (SEMA A), full insert sequence.
AK043634
ACCESSION
VERSION AK043634.1 GI:26089859
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus
ORGANISM
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED 10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
REFERENCE
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staubli, P., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Machionni, L., Maehima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilm, D.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

21085660
11217851

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 2960)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

1. .2960
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:A830014F15"
/db_xref="taxon:10090"
/clone="A830014F15"
/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
294. .2543

misc_feature

/note="SEMAPHORIN 3B PRECURSOR (SEMAPHORIN A) (SEMA A)
(SWISSPROT|Q62177, evidence: FASTA, 98.3%ID, 100%length,
match=2247)
putative"

polya_signal

/note="putative"

polya_site

/note="putative"

BASE COUNT 624 a 837 c 874 g 625 t

ORIGIN

Query Match 22.9%; Score 620.2; DB 11; Length 2960;
Best Local Similarity 59.5%; Pred. No. 2.4e-146; Indels 21; Gaps 4;
Matches 1130; Conservative 0; Mismatches 748;
QY 359 AGCTCCAGTATATCATCTTCTTCTTGGATGAGGAAACGAGTAGCTCTATGTTGGAGCA 418
Db 447 ACCTGCTGTTATGAGCGTCTGCTGTTGATGAGGAGCGTGACCGCTGTTTGGTGTCT 506
QY 419 AAGGATCACATATTTTCATTCGACCTGGTTAATAT---CAAGATTTTCAAAAGATTGTG 475
Db 507 GAGAACCAAGTGGGCTTCCCTCAGCTGAGCAACATCAGCAAGCGGCAAGAGCTGGCC 566
QY 476 TGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGGGCTGGAAGAAGACATCTTGAA 535
Db 567 TGGCCCGCCCGCTGGAATGGCGTGAAGATGCAACTGGGCGAGGAAGCAATGGGTACC 626
QY 536 GAATGTGCTAATTTTCATCAAGGTACTTAAGGCATATATCAGACTCACTTGTAGCCCTGT 595
Db 627 GAGTGCTATGAACTTCGTGAGCTGCTGCACACCTACACACACCCACTTGTGGCTGT 686
QY 596 GGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTTGAAATTTGACATCATCTGAGGAC 655
Db 687 GGCACAGGGGCTTCCACCCCAACCTGTGCTTTTGTGAGGTGGGCCACCGCTGGAGGA 746
QY 656 AATATTTTAACTGGGAGAACTCACATTTTGAAGACCGCGTGGGAGAGTCCATATGAC 715
Db 747 CCCATGCTTCACTGGACCGGAGGAACTTGGAGACCGGCAAGGGGAGAGACTCTTTATGAC 806
QY 716 CCTAAGCTGTGCAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAATCTGCAGCT 775
Db 807 CCAAGGATCGGGCTGCTGCTGGTGGGGAAGAACTGTATTTCTGGGGTGACAGCA 866
QY 776 GATTTTATGGGGGAGACTTTGCTATCTTCCGAATCTTGGGACACCAACCCCAATCAGG 835
Db 867 GACCTTATGGGCGGGACTTTTACCCTCTTTGGAAGCCTTGTGTCAGAAATCCGAGTCTCGA 926
QY 836 ACAGAGCAGATGATTTCCAGTGGCTCAATGATCCAAGTTCATTAGTGGCCACTCATC 895
Db 927 ACAGAGCCCATGATTTCCCGCTGGCTCAATGAACCCCAAGTTTGTCAAGGCTCTTTGGATC 986
QY 896 TCAGAGAGTGACAATCTCTGAAGATGACAAAGTATACATTTTCTCCGTGAAAATGCAATA 955
Db 987 CCAGAGAGTGAGAACCTTGATGACGATAAATCTATTCTTCTCCGAGTCCGCTGTG 1046
QY 956 GATGAGAAACACTC---TGAAAAGCTTACTCACGCTAGATAGTATGATGATGCAAGAAT 1012
Db 1047 GAAGCAGCACAGCAATGGGGCGCATGCTGTGCTGCTGTGTTGGCCAGATCTGCAGGAAT 1106
QY 1013 GACTTTGGAGGGCACAGAGTCTGGTGAATAAATGGACACACTTCTCAAGCTCGTCTG 1072
Db 1107 GACCTGGTGGCCAGCGGAGCTTGGTCAACAAATGGACCACTTTCTGAAGGCGCGGCTT 1166
QY 1073 ATTTGCTCAGTGGCCAGGTCCAAATGGCATTTGACACTCATTTTGTAGTAAGTGCAGATGTA 1132
Db 1167 GTGTGCTCAGTACCTGG---AGTTGAGGGTGACACCCACTTTGACCAACTTCAGAGTGT 1223
QY 1133 TTCTTAATGAACCTTTAAAGATCTTAAATAATCCAGTGTGATATGAGAGTGTGTACGATTC 1192
Db 1224 TTCTTCTGCTCTCCGAGACCGCCAGACACCTTCTCTATGCTGTCTCTCCACCTCC 1283
QY 1193 AGTAACATTTTCAAGGATCAGCGGTGTGTATGTTATAGCATGATGATGATGAGAGGGTGTG 1252
Db 1284 AGTGGTGTCTTCCAGGGCTCTGTGTGTGCTGTGATGAGCATGATGATGATGAGAGGGTGTG 1343
QY 1253 TTCTTGGTCCATATGCCACAGGATGAGCCCAACTATCAATGGGTGGCTTTATCAAGGA 1312
Db 1344 TTCTTGGACCTTTTGTCTCAAAAGAGGGGCTTACACACAGTGGGTGCTCTACAGGGT 1403
QY 1313 AGAGTCCCTTATCCAGGGCAGGAATTTGTGCCAGCAAAACATTTGGTGGTGTGTTGACTCT 1372
Db 1404 CGTGTCCCTTACCCAGAGACCTTGTCATGTGCCCCAGCAAGACCTTTGGCACCTTTCAGCTCC 1463

QY	1373	ACAAGGACCTTCTCTGATGATGTATTAAACTTTGGCAAGAAGTCATCAGCCATGTACAAT	1433
Db	1464	ACCAAGGACTTCCAGATGACGTTTATCCAGTTTGCTCGGAACACCCTCTCATGTACAAC	1523
QY	1433	CCAGTGTTCCTATGAACAATCGCCCAATAGTAGTCAAACCGATGTAAATTTATCAATTT	1492
Db	1524	CCAGTCTGCCATGGGGGGCGCCCTCTCTTCTTAAGTGGAGCTGGGTACACCTTC	1583
QY	1493	ACACAAATTTCTGTAGAACCGAGTGGATGCAGAGAAGATGGACAGTATGATGTATTGTTTTATC	1552
Db	1584	ACCCAAATCGCGCAGACCGAGTAGCAGCTGCCGATGACACTACGATGTTCTCTTCATTT	1643
QY	1553	GGAACAGATGTTGGGACCGTTCTTAAAGTAGTTTTCATTCCTAAGAGAGACTTCGTATGAT	1612
Db	1644	GGTACAGATGTGGCACAGTGTCTGAAAGTAGATCTCAGTCCCCTAAAGACGACCGACCTAAT	1703
QY	1613	TTAGAAGAGTGTCTGTGGAAGAAATGACAGTTTTTTCGGGAACCGACTGTATTTCAGCA	1672
Db	1704	TCTGAAGGACTTCTCTGGAAGAGCTGCAGTGTTCGAGAGCTCTGCCGTATACACGC	1763
QY	1673	ATGAGGCTTTCCTAAGCAGCAACAATAATATTGGTTTCAACGGCTGGGGTTGGCCAG	1732
Db	1764	ATGCAAAATCTCTTAAAGGCAACAACCTCTACATAGCATCGCGCAGCGAGTGGCCAG	1823
QY	1733	CTCCCTTTACACGGTGTGATATTACGGGAAGCGTGTCTCAGTGTTCCTCGGCCGA	1792
Db	1824	ATTGCTTTGTCATCGCTGCATCGCCTAGGCGCGCCCTCGCAGCAATGCTCTTTGGCCGT	1883
QY	1793	GACCTTACTGTGCTTGGGATGGTTCTGCAATGTTCTCGCTATTTTCCACTGCAAAAGAGA	1852
Db	1884	GATCTTACTGCGCTGGGATGATCAGCTTGCACAGCTTCAGCCTCAGGCTCAGGCAAGAGA	1943
QY	1853	CGCACAGACGACAAGATATPAGAAATGGAGACCCACTGACTCACTGTTTCAGACTTTACAC	1912
Db	1944	CGGTTCCGGAGGCAAGACATAAAGAAATGGCGACCCCGACGACCCCTATGCTC-----	1993
QY	1913	CATGATAATCACCATTGSCCAGCCCTGAGAGAGATCATCTATGGTGTAGAGAATAGT	1972
Db	1994	--TGGAGACTCTTCTCACTCTGTGCTGTGGAGAAAGGTGTGGGTGTGGAGAGCGCG	2051
QY	1973	AGCACATTTTGGAAATGCAGTCCGAAAGTCGACAGAGCGCTGTCTATTGGCAATTCAG	2032
Db	2052	AGCGCGTTTCTGGAGTGTGAGCCCGCTCGCTCCAGGCGCATGTGCAGTGGACTTCCA	2111
QY	2033	AGCGGAATGAAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATPCAGACAGAT	2092
Db	2112	GGTGCAGGGGAGCGAGCTCACACCCAGGTCTGSGCTCAGGAGAGAGTAGAGCGCACTGG	2171
QY	2093	CAAGGCTTCTGCTAGTACTCTACAACAGAGGATTCAGGCAATTAACCTCTGCCATGG	2152
Db	2172	CGGGGCTGCTGTGTGGGGGGCTGCGGCGCGAGACTCTGGCGTGTATCTTTGCGTCGG	2231
QY	2153	GTGGAACATGGGTTTCATACAAACTCTTCTTAAAGTAAACCCCTGGAAGTCAATTGACAGAG	2212
Db	2232	GTTGAAACAGGCTTTTCACAAACCACTGGCTGGCTGGTGTGCTATGTTGTAGTGGCGG	2291
QY	2213	CATTGGAAGAACTTCTTCAATAAGATGATGAGAGAT	2251
Db	2292	CAGGCTGAACGACTGGGCACGGGCAGAGGAAGCAGCGCT	2330

RESULT 8
AK033597

LOCUS AK033597 3154 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:913002F09 product:SEMAPHORIN 3B PRECURSOR (SEMAPHORIN A) (SEMA A), full insert sequence.

ACCESSION AK033597

VERSION AK033597.i GI:26083479

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1	
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
2	
AUTHORS	Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
3	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Katsuno, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
4	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fletschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, I., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
5	
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	(bases 1 to 3154)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission

QY 1913 CATGATATACCAATGCGCAGAGGAGAGATCATCTATGTTGAGAGATAGT 1972
 D5 --TGGAGACTCTTCTCACTCTGCTGCTGAGAGAGAGAGGTTGGGTGGAGCGGC 2108
 QY 1973 AGCAATTTTGGATGAGTCCGAGTCCGAGAGAGGCTGCTATGTCATTCAG 2032
 D5 AGCGGTTTCTGGAGTGTGAGCCCGCTGCTCCAGCGCATGTGCACTGAGACCTCCAA 2168
 QY 2033 AGGCGAATGAGAGAGAGAGAGAGATCAGATGAGATGATCATATCATCAGACAGAT 2092
 D5 GTGAGTGGAGGAGCTCACACAGTGTGCTGAGGAGAGATGAGAGGAGCTGCG 2228
 QY 2093 CAAGCCTTCTGCTAGTGTACAGAGAGAGATTCAGCAATTCCTCTGCACTGCG 2152
 D5 CGGGGCTGCTGTTGCGGGGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2288
 QY 2153 GTGAGATGAGTTCATCAAACTCTTCTTAAGTAACTGAGGAGGAGGAGGAGGAG 2212
 D5 GTTGAACAGAGCTTTTCAACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2348
 QY 2213 CATTTGGAAGAACTTCTTCTTATAGATGATGATGAGAT 2251
 D5 CAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2387
 ACCESSION AK037034 3292 bp mRNA linear HTC 05-DEC-2002
 LOCUS Mus musculus adult female vagina cDNA, RIKEN full-length enriched
 DEFINITION (SEMAPHORIN A) (SEMA A), full insert sequence.
 ACCESSION AK037034
 VERSION AK037034.1 GI:26085567
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazawa, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, S.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staab, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaj, H., Kontecki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 3292)
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohata, N.,
 Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 FEATURES
 Location/Qualifiers
 1. 3292
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:9930101102"
 /db_xref="taxon:10090"
 /clone="9930101102"
 /sex="female"
 /tissue types="vagina"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 misc_feature 389..2638
 /note="SEMAPHORIN 3B PRECURSOR (SEMAPHORIN A) (SEMA A)
 (SWISSPROT|Q62177, evidence: FASTA, 98.3%ID, 100%length,
 match=2247)
 putative"
 BASE COUNT 699 a 922 c 967 g 704 t
 ORIGIN
 Query Match 22.9%; Score 620.2; DB 11; Length 3292;

AUTHORS

Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seva, I., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.

FUNCTIONAL annotation of a full-length mouse cDNA collection

Nature 409 (6821), 565-590 (2001)

21085660

11217851

REFERENCE

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 12127851
 PUBMED 12127851

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Atakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, I.,
 Sogabe, Y., Takeda, Y., Tanaka, T., Tomaru, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

Yokohama 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

COMMENT

FEATURES

source

misc_feature

polyA_signal

polyA_site

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1128; Conservative

Score 617; DB 11; Length 3287;

Pred. No. 1.6e-145;

Mismatches 750; Indels 21; Gaps 4;

359 AGCTCCAGTATCATACCTCTCTTTGGATCAGAACGAGCTGCTGTATTTGGAGCA 418

775 ACCTGCTGTATGAGCGCTTCTGCTGCTGAGGAGCGTGGACGCTTGTGGTGCT 834

419 AAGGATCACATATTTTCATTCGACCTGGTGTATATAT---CAAGGATTTTCAAAAGATTGTG 475

835 GAGAACACGCTGGCTTCCCTCAGCTGGACCAACATCAGCAAGCGAGCGAGCTGCC 894

476 TGGCCAGTATCTTACACGAGAGAGATGATGCAAGTGGGCTGGAAAAGACATCTCGAAA 535

895 TGGCCCGCCCGCTGGGATGCGGTGAAGATGCAACTGGGCGAGGAGCATTTGGTACC 954

536 GAATGTGCTAATTTTCATCAAGGTACTTAAAGCATATATATCAGCTCCTTGTAGCCCTGT 595

955 GAGTGATGATGACTTCTGTAAGCTGCTGACACCTACCAACACACCTCTGCTGCTGCTGT 1014

596 GGAACGGGGGCTTTTCATCCAAATTTGACCTTACATTTGAAATTTGACATCATCTTGAGGAC 655

1015 GGCACAGGGGCTTTCCACCCACCTGTCCTTTTGGAGGTTGGCCCGGCTGGAGGAA 1074

656 AATATTTTAAAGCTGGAGAACTCATTTTGAACCGGCGCTGGAGAGAGTCCATATGAC 715

1075 CCCATGCTTCACTGGAGCGGAGAACTTGGAGCGGAGGAGGAGAGACTCTTATGAC 1134

716 CCTAAGTCTGACAGCATCCCTTTTAAATAGATGAGATTTATCTCTGGAAGTCCAGCT 775

1135 CCAAGGATCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194

776 GATTTTATGGGGGAGAGCTTTGCTATCTTCGAACTCTTGGGACCAACCAACCAACCAAG 835

1195 GACCTTATGGCGGGGACTTTACCATCTTTTGAAGCTTTGGTCAAGATCCGAGTCTCGA 1254

836 ACAGAGCAGCATGATCCAGGTGGCTCAATGATCCAAAGTTTATTAGTGGCCCACTCATC 895

```
Db 1255 ACAGAGCCCCATGATTCGCGCTGGCTCAATGAACCAAGTTTGTCAAGGTCTTTGGATC 1314
Qy 896 TCAGAGAGTGAACAATCCTCAAGATGACAAAGTATATCTTTTCTCGTGGAAATGCAATA 955
Db 1315 CCAGAGAGTGAGAACCTTGATGACGATAAATATCTATTTCTTCCGAGTCCGCTGTG 1374
Qy 956 GATGAGAACACTC---TGAAAAAGCTACTCACGCTAGATAATAGGTGAGATATGCAAGAT 1012
Db 1375 GAAGCAGACACCAAGCAATGGCGCATGCTGTGTCTCTGTTGGCCAGATCTGCAAGAT 1434
Qy 1013 GACTTTGAGGGCAGCAAGAGTCTGTGTAATAATGGAACAATCTCTCAAGCTGCTGTG 1072
Db 1435 GACCTGGGTGGCGAGCGGAGCTTGTGTCAAAATGGACCAATTTCTGAAGCGGGGTT 1494
Qy 1073 ATTTGCTCAGTGCAGGTCCTCAATGCAATGCACTCAATTTTGTGATGAACTGCAAGATGA 1132
Db 1495 GTGTGCTCAGTACCTGG---AGTTGAGGTGACACCCACTTTGACCAACTTCAGGATGT 1551
Qy 1133 TTCTTAATGAATCTTAAAGATCCTTAAAGATCCAGTTGTATATGAGAGTGTTTACGACTTCC 1192
Db 1552 TTCTCTGTCTCTCCGAGACCCGACACACCTCTTCTATGCTGTCTTCTCCACTCC 1611
Qy 1193 AGTAACATTTTCAAGGATCAGCGTGTGTATGTATAGCATGAGTGTGTGAGAAAGGTG 1252
Db 1612 AGTGTGTCTTCCAGGGCTCTGCTGTGTGTGTATACAGCATGAACGATGTGCGCGAGCC 1671
Qy 1253 TTCTTTGTTGTCATATGCCACAGGATGACCACTATCAATGATGGTGCCTTATCAAGA 1312
Db 1672 TTCTTTGGAGCTTTTGTCTCACAAGAGGGGCTTACACACAGTGGGTGTCTTACCAGGT 1731
Qy 1313 AGAGTCCCTATCCAGCGCCAGGAATTTGTCCAGCAAAATTTTGGTGGTTTGACTCT 1372
Db 1732 CGTGTCCCTACCCAGACCTGCGATGTGCCCCAGCAAGACCTTTGGCACCTTCAGCTCC 1791
Qy 1373 ACAAGAGACCTTCTGTATGATGTTATTAACCTTTTGAAGAAGTCAATCCAGCCATGACAT 1432
Db 1792 ACCAAGGACTTCCAGATGACGTTTATCAGTTTGTGCGAACCACCTCTCATGTACAA 1851
Qy 1433 CCAGTGTTCCTATCAACAATCGCCCAATAGTATCAAAACGATGAATTTATCAATTT 1492
Db 1852 CAGTCTTCCCATGCGGGGGCGCCCTCTTCTTCAAGTGGGAGCTGGGTACACCTTC 1911
Qy 1493 ACACAAATTTGCTAGACCGAGTGTGATGAGAGATGACATGATGATGTTATGTTTATC 1552
Db 1912 ACCAAATCGCGCAGACCGAGTAGCAGTGCAGTGCAGTACGATGATGTTCTTCAAT 1971
Qy 1553 GGAACAGATGTTGGACCGCTTCTTAAAGTATGTTTCAATCTTAAAGGAGACTTGTGATGAT 1612
Db 1972 GGTACAGATGTGGGCACAGTGTGAAGTGTATCTCAGTCCCCCAAGGCGAGCCACTAAT 2031
Qy 1613 TTAGAAGAGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACGACTGCTATTTCAGCA 1672
Db 2032 TCTGAAGGACTTCTCTGGAAGAGCTGAGGTTTCGAGGACTCTGCGCTATCACAGC 2091
Qy 1673 ATGAGACTTTTCCACTAAGCAGCAACAATATATATTTGTTTCAACGCTGGGGTGGCCAG 1732
Db 2092 ATGCAATCTCTCTTAAAGGCAACAATCTACATAGCATCGCGAGCGAGTGGCCAG 2151
Qy 1733 CTCCCTTTACCGGTGTGATATTTACGGGAAGCGTGTGTGAGTGTGCTCGCCCGA 1792
Db 2152 ATTTGTTTGCATCGCTGCACTGCCCTAGCGCGCGCTCGCGAGATGCTGTGTCGCCGT 2211
Qy 1793 GACCTTACTGTCTTGGATGGTTCTCAGTGTCTCGCTATTTTCCCACTGCAAGAGA 1852
Db 2212 GATCTTACTAGCTTGGATGGATCAGCTTGCAACGCTTCCAGCCTACGCGCAAGAGA 2271
Qy 1853 CGCAACAGCAGCAGATATAGAATAGGAGACCCACTCACTGCTTTCAGACTTACAC 1912
Db 2272 CGTTCCGAGGCAAGACATAAGGATGGCGACCCCGACACCTATGCTC----- 2321
Qy 1913 CATGATATACCATGGCCACAGCCCTGAAGAGAGAAATCATCTATGTTGTGAGATAAGT 1972
Db 2322 ---TGAGACTCTTCTCACTCTGTGCTGTGGAGAGAGAGGTGTTGGGTGTGGAGAGCGC 2379
```

```
Qy 1973 AGCATTTTTGGATTCAGTCCGAAGTCCGAGAGCGCTGTCTATTGGCAATTCAG 2032
Db 2380 AGCGGTTTCTGGAGTGTGAGCCCGCTCGTCCAGGGCGCATGTGCACTGGACCTTCCA 2439
Qy 2033 AGGCAATTAAGAGAGCGAAAAAGAGATCAGAGTGGATGATCATATCATCAGACAGAT 2092
Db 2440 GGTGAGGGAGGAGGAGCTCACACCAGGTGCTGGCTGAGGAGAGATGAGCGCACTGGC 2499
Qy 2093 CAAGCCCTTCTGCTACGTAGTCTACACAGAAAGATTCAGGCAATTAATCTCTGCGCATGG 2152
Db 2500 CGGGGCTGCTGTGTTGCGGGGCTGCGGCGCAGGACTCTGCGGTGTATCTTTTGGTGGCG 2559
Qy 2153 GTGGAACATGTTTTCATACAACTCTTCTTAAGGTAACCTTGAAGTCAATTCAGACAGAG 2212
Db 2560 GTTGAACAAGACTTTTCAACCACTGCTGCTGCGCTGCTGTCATGTGTTGAGTGGCG 2619
Qy 2213 CATTTGGAAGAACTTCTTTATAAAGATGATGAGAT 2251
Db 2620 CAGGCTGAACGACTGGCAAGGCGGAGAGGAGCGCGCT 2658
```

RESULT 11

```
CD349263
LOCUS
DEFINITION
IMAGE: 6854395 5', mRNA sequence.
ACCESSION
VERSION
CD349263.1 GI:31140778
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 729)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Straubeberg, Ph.D.
Email: c8apbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
```

Seq primer: pyX-5.

FEATURES

source

Location/Qualifiers

```
1..729
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6854395"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP Fv0"
/Note="Organ: Brain; Vector: pyX- Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
```

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 210 a 165 c 160 g 194 t

Query Match 21.9%; Score 594.6; DB 14; Length 729;
Best Local Similarity 88.5%; Pred. No. 4.4e-140;
Matches 645; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 1032 GTCTGGTGAATATAATGGAACAATCCCTCAAGCTCGCTGATTTGCTCAGTGCAGTC 1091
Db 1 GTCTGTGAATATAATGGAACAATCCCTCAAGCTCGCTGATTTGCTCAGTGCAGTC 60

Qy 1092 CAATGGCAATGACACTCATTTTGAATGAATGCGAGATGTATTCCTAATGAACCTTTAAAG 1151
Db 61 CCAATGGCAATGACACTCATTTTGAATGAATGCGAGATGTATTCCTAATGAACCTTTAAAG 120

Qy 1152 ATCTTAAATCCAGTTGTATGAGGTGTTTACGACTTCCAGTAAACATTTTCAAGGAT 1211
Db 121 ATCTTAAATCCAGTTGTATGAGGTGTTTACGACTTCCAGTAAACATTTTCAAGGAT 180

Qy 1212 CAGCGTGTATGTATAGCATGATGTAGAGAGGTGTTCTTGGTCCATATGCC 1271
Db 181 CTGCTGTGTATGTATAGCATGATGTAGAGAGGTGTTCTTGGTCCATATGCC 240

Qy 1272 ACAGGATGGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGC 1331
Db 241 ACAGGATGGTCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGC 300

Qy 1332 CAGGAATCTGTCGAGCAAAACATTTGGTGTGTTGACTCTCAAGGACCTTCTGATG 1391
Db 301 CAGGAATCTGTCGAGCAAAACATTTGGTGTGTTGACTCTCAAGGACCTTCTGATG 360

Qy 1392 ATGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAATCCAGTGTTCCTATGAACA 1451
Db 361 ATGTCATAACCTTTGCAAGAGTCAATCCAGCATGTACAATCCAGTGTTCCTATGAACA 420

Qy 1452 ATCCGCCAATAGTATCAAAACGAGTGTAAATTTATCAAAATTTGCTGAGACC 1511
Db 421 ATCCGCCAATAGTATCAAAACGAGTGTAAATTTATCAAAATTTGCTGAGACC 480

Qy 1512 GAGTGGATGCGAGAGATGCGAGTATGATTTATCGGAACAGATGTTGGGACCG 1571
Db 481 GAGTGGATGCGAGAGATGCGGAGTATGATTTATCGGAACAGATGTTGGGACCG 540

Qy 1572 TTCTTAAAGTGGTTTCAATTCCTAAGGAGACTTGTATGATTTAGAGAGTCTGCTGG 1631
Db 541 TTCTTAAAGTGGTTTCAATTCCTAAGGAGACTTGTATGATTTAGAGAGTCTTCTTGG 600

Qy 1632 AAGAAATGACAGTGTTCGGAACCGACTGCTATTTCAAGCAATGAGGCTTTCCACTAAGC 1691
Db 601 AAGAAATGACAGTGTTCGGAACCGACTGCTATTTCAAGCAATGAGGCTTTCTACTAAGC 660

Qy 1692 AGCAACACTATATATTTGGTTCACAGGTGCGGTTGCCAGTCCCTTTTACCGGTTGT 1751
Db 661 AGCAACACTGTATATTTGGTTCACAGGTGCGGTTGCCAGTCCCTTTTACCGGTTGT 720

Qy 1752 ATATTTAGC 1760
Db 721 ACACTATG 729

RESULT 12
CB196368
LOCUS CB196368 900 bp mRNA linear EST 05-FEB-2003
DEFINITION AGNCOURT_11259678 NIH_MGC_135 Mus musculus cDNA clone.
IMAGE:30138167 5', mRNA sequence.
ACCESSION CB196368
VERSION CB196368.1 GI:28224005
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 900)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0040 row: i column: 24
High quality sequence stop: 655.

FEATURES

source

1..900
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clones="IMAGE:30138167"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/notes="vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5
, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTAGTTCTAGATCGAGCGCGCCCTT-3' Tissue contributed by
David Rowe. Library constructed by ResGen, Invitrogen
Corp."

BASE COUNT 286 a 208 c 225 g 178 t 3 others
ORIGIN

Query Match 21.8%; Score 590.4; DB 14; Length 900;
Best Local Similarity 83.0%; Pred. No. 5.6e-139;
Matches 747; Conservative 0; Mismatches 138; Indels 15; Gaps 6;

Qy 1700 CTATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGATATTTAC 1759
Db 1 CTGTACATTTGGCTCAACTGCGGGAGTGGCAAGCTTCTCTACACCGCTGTGACATCTAT 60

Qy 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGTGGATGGTCT 1819
Db 61 GGCAAGCGCTGTGAGAATGCTGCTCGTGGGACCTTACTGTGCTGGGTGGGTCC 120

Qy 1820 GCATGTTCTCGCTATTTTCCACCTGCAAGAGACGCAAGACCAAGATATAAGAAAT 1879
Db 121 TCATGCTCACGCTATTTTCTTACTGCAAGAGGCGCACAGCCAGATATAAGGAAT 180

Qy 1880 GGAGACCACTGACTCACTGTTCAGACTT---ACACCATGATATCACCATGGCCACAGC 1936
Db 181 GGAGACCACTGACTCACTGTTCAGACTTGGAGCACCATGATATCATCTGGGCCACG 240

Qy 1937 CCTGAAGAGAGATCATCTATGTTAGAGATAGTAGCATTCTTGGATGCGATCCG 1996
Db 241 CTTGAAGAGAGATCATCTATGAGTGGAAACAGTAGTACATTTCTTGGATGCGATCCG 300

Qy 1997 AAGTCGAGAGAGCGCTGGTCTATTGGCAATTCAGAGGGGAAATGAAGCGCAAAAGAA 2056
Db 301 AAGTCAGAGAGCGCTGGTATATTGGCAATTCAGAGGAGAAATGAAGTCGAAGAG 360

Qy 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTA 2116
Db 361 GAGATCAGAGTGGTGTATCATATCATCAGGACAGAAACAAAGGCTCTCTGCTCCGTAGCCTG 420

Qy 2117 CAACAGAGGATTCAGGCAATTTACCTCTGCCATCGGTTGGACATGGTTCATACAACT 2176
Db 421 CAGAAAGAGGATTCAGGCAATTTACCTCTGTGCACTGTGGAAACACGGATTTCATCAAACT 480

QY 2177 CTTCTTAAGGTAAACCTGGAAGTCAATTGACACAGAGCAATTTGGAGAAGCTTCTTCATATA 2236
 Db 481 CTTCTTAAGGTAAACCTGGAAGTCAATTGACACAGAGCAATTTGGAGAAGCTTCTTCATATA 540
 QY 2237 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCAATAGATGACACCTAGCCAG 2296
 Db 541 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCAATAGATGACACCTAGCCAG 600
 QY 2297 AAGGTCTGTGTACAGAGATTTCAATGAGCTCATCAACCCCAATCTCAACACAGATGGAT 2356
 Db 601 AAGGTCTGTGTACAGAGATTTCAATGAGCTCATCAACCCCAATCTCAACACAGATGGAT 660
 QY 2357 GAGTCTGTGTAAACAAGTTTGGAAAAGGACCGAAAACCAAGCTCGGCAAGGCGGAGCAT 2416
 Db 661 GAGTCTGTGTAAACAAGTTTGGAAAAGGACCGAAAACCAAGCTCGGCAAGGCGGAGCAT 719
 QY 2417 ACCCAGGGAACAGTAACAATGGAAGACCTTACAGAAATTAAGAAAGGTAGAAAC-AG 2475
 Db 720 TCTCANGGGAGCAGCAACAGTGGAGGACATGCCAGAGAGCAAGAAAGGTAGAAACNAG 779
 QY 2476 GAGGACCCACGAATTT--GAGAGGGCACCAGGAGTGTCTGA---GCTGCATTACCTCT 2529
 Db 780 GAGGACCCACGAATTTGAGCGGGGACCCAAATGTCTGAAGCTGGCCCACTTCCA 839
 QY 2530 AGAAACCTCAAAACAAGTAGAAACCTTGCCCTAGACAATAA----CTGGAATAAACAATGCAA 2585
 Db 840 AAACCTCAAAACAAGTAGAAACCTTGCTTTAGATAATAACCTTGGAATAAACAATGCAA 899

RESULT 13

BF667677 889 bp mRNA linear EST 21-DEC-2000
 LOCUS 602121773F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810 5',
 DEFINITION mRNA sequence.
 ACCESSION BF667677.1 GI:11941572
 VERSION EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCm1103 row: a column: 19
 High quality sequence stop: 536.
 Location/Qualifiers
 1..889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4278810"
 /tissue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_56"
 /note="Organ: Brain; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggccgctggcc); Site 2: SfiI (ggccattatggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor
 sequence: 5'-ATTTAGAGCCGAGGCGGCGGAGCATG-dr(30)BN-3'
 (where B = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies

FEATURES

SOURCE

RESULT 14
 BU703431

LOCUS

DEFINITION

IMAGE: 6405252 5', mRNA sequence.

ACCESSION

BU703431

VERSION

BU703431.1

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 775)

AUTHORS

NIH-MGC <http://mgi.nci.nih.gov/>.

contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA).
 BASE COUNT 263 a 158 c 250 g 217 t 1 others

ORIGIN

Query Match 21.7%; Score 588.6; DB 10; Length 889;

Best Local Similarity 95.9%; Pred. No. 1.6e-138;

Matches 647; Conservative 0; Mismatches 24; Indels 4; Gaps 4;

QY 1595 AAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGAAGAAATGACAGTTTTTCGGAA 1654

Db 1 AAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGAAGAAATGACAGTTTTTCGGAA 60

QY 1655 CCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCAGCAACTATATATATGTTCA 1714

Db 61 CCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCAGCAACTATATATATGTTCA 119

QY 1715 ACGCTGGGGTTGCCAGCTCCCTTTACACGGTGTGATATTTACGGGAAGCGTGTCT 1774

Db 120 ACGCTGGGGTTGCCAGCTCCCTTTACACGGTGTGATATTTACGGGAAGCGTGTCT 179

QY 1775 GAGTGTGCTCGCCGAGACCTTACTGTCTTGGATGCTTCTGCAATGTTCTGCTAT 1834

Db 180 GAGTGTGCTCGCCGAGACCTTACTGTCTTGGATGCTTCTGCAATGTTCTGCTAT 239

QY 1835 TTTCCCACTGCAAGAGACGCAAGAGCAAGATATAAGAAATGGAGACCCACTGACT 1894

Db 240 TTTCCCACTGCAAGAGACGCAAGAGCAAGATATAAGAAATGGAGACCCACTGACT 299

QY 1895 CACTGTTTACAGCTTACACCATGATTAATCACCATGCGCCACAGCCCTGGAAGAGAGATCATC 1954

Db 300 CACTGTTTACAGCTTACACCATGATTAATCACCATGCGCCACAGCCCTGGAAGAGAGATCATC 359

QY 1955 TATGCTGATGAGATAGTACACATTTTGGAAATGCAAGTCCGAGTCCGAGAGCGGTG 2014

Db 360 TATGCTGATGAGATAGTACACATTTTGGAAATGCAAGTCCGAGTCCGAGAGCGGTG 419

QY 2015 GTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGATCAGAGTGGATGAT 2074

Db 420 GTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGATCAGAGTGGATGAT 479

QY 2075 CATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTTACCAAGAGAGATTCAGGC 2134

Db 480 CATATCATCAGGACAGATCAAGG-CTTCTGCTAGTGTCTTACCAAGAGAGATTCAGGC 538

QY 2135 AATTACCTCTGCTGCTGAGATGAGTGGTTCATCAAACTCTTCTTAAGTAACTCTG 2194

Db 539 AATTACCTCTG-CTGCGGTGGAAATGAGTGGTTCATCAAACTCTTCTTAAGTAACTCTG 597

QY 2195 GAAGTCAATTCAGACAGAGCATTTTGAAGAACTTCTTCAATAAGATGATGAGAGATGGC 2254

Db 598 GAAGTCA-CTGGCAGAGCTTTGGGAAGAAATTTCTTCTTAAGATGATGAGAGATGGC 656

QY 2255 TCTAAGACCAAGAA 2269

Db 657 TCTAAGACCAAGAA 671

775 bp mRNA linear EST 09-OCT-2002

UI-M-FOO-bzo-k-13-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone

IMAGE: 6405252 5', mRNA sequence.

ACCESSION

BU703431

VERSION

BU703431.1

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 775)

AUTHORS

NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyX-5.
 Location/Qualifiers
 1. 699
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6405252"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_F00"
 /note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TCAGAGACC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
 221 a 184 c 165 g 200 t 5 others

Query Match 21.4%; Score 581; DB 13; Length 775;
 Best Local Similarity 86.5%; Pred. NO. 1.3e-136;
 Matches 661; Conservative 0; Mismatches 100; Indels 3; Gaps 2;
 708 CATATGACCTTAAGTGTGACAGATCCCTTTTAATAGATGGAGATTAATCTGGAA 767
 1 CTTATGATCCCAAACTACTGACGCTCTCTCTTAATAGACGGTGAGTGTCTGGAA 60
 768 CTGACGCTGATTTATGGGGGAGACTTTGCTATCTCCGAACCTTGGGCACCAACCAC 827
 61 CTGCTGCGGACTCATGGGAGCGGACTTCGCTATCTTCAGAACACTGGGGCACCATCACC 120
 828 CAATCAGACAGACAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAATAGTGCCC 887
 121 CCATCAGACGAGCAGCATGACTCCCGTGGCTCAATGATCCTAGATTCATAGTGCCC 180
 888 ACCTCATCTCAGAGTGACAATCTCTGAAGATGACAAAGTATACATTTTCTTCGGTGAAA 947
 181 ATCTCATCCAGAGTGACAACTCTGAAGATGACAAAGTATATTTTCTTCGGAGAAA 240
 948 ATGCAATAGATGGAGAACACTCTGGAAAGCTACTCAGCTAGATAGTGCAGATATGCA 1007
 241 ATGCAATAGACGGAGAACACTCTGGAAAGCACTCATGCTAGATAGTGCAGATATGCA 300
 1008 AGAATGACTTTGGAGGGCAGCAAGTCTGGTGAATTAATGACACAACTTCCTCAAAGCTC 1067
 301 AGAATGACTTTGGTGGACAGAGAGTCTTGTGAATTAATGACACAACTTCCTCAAAGCTC 360
 1068 GTCTGATTTGCTAGTGCCAGGTCCAAATGGCACTCATTTTGATGAACCTGACGG 1127

Db 361 GCCTGATTGCTGTGTGCCCGGTCCCAATAGGCATGTGACCCATTTTGATGAATTGCAGG 420
 Qy 1128 ATGTATTCTTAATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATATGAGTGTTCAGG 1187
 Db 421 ATGTATTCTTAATGAACCTTTAAAGATCCTAAAAATCCAGTGTCTATGAGTGTTCACAA 480
 Qy 1188 CTTCCAGTAACATTTTCAAGGGATCAGCCGTGTGTATGTATAGCATGATGTGAGAA 1247
 Db 481 CATCAAGCAACATCTTTAAAGGATCTGTGTGTGCATGTACAGCATGATGTGATGAAGA 540
 Qy 1248 GGGTGTCTTGTGTCATATGCCCCACAGGGATGACCACTATCAATGGGTGCTTATC 1307
 Db 541 GGGTGTCTTGTGTCATATGCTCACAGAGATGGTCCCACTATCAGTGGGTGCTTACC 600
 Qy 1308 AAGGAAGAGTCCCTTATCCAGGGCAGGAACCTTGTCCAGCAAAACATTTGGTGGTTTTG 1367
 Db 601 AAGGAAGAGTCCCTTATCCAGGGCAGGAACCTTGTCCAGTAAAAACATTTGGCGGATTTG 660
 Qy 1368 ACTCTACAAGGACCTTCTCTGATGATGTTATTAACCTTTGCAAGAGTCATCCAGCCATGT 1427
 Db 661 ACTCCACAAGGACCTTCTCTGATGATGTCATACCTTTTGCAGAA-TCATCCAGCCATGT 719
 Qy 1428 ACATCCAGTGTCTTCTATGAACAATGCCCAATAGTATGATCAAA 1471
 Db 720 ACAACAGTGTTCCTAT--AATATCGCCGATCATGATCANA 761

RESULT 15
 CA327319 699 bp mRNA linear EST 27-NOV-2002
 LOCUS UI-M-FY0-ccy-p-01-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE: 6826322 5', mRNA sequence.
 CA327319
 ACCESSION CA327319.1 GI:24545417
 VERSION EST.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 699)
 NTH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyX-5.
 Location/Qualifiers
 1. 699
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_lib="NIH_BMAP_FY0"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chan, Ph.D., program coordinator."

BASE COUNT	197 a	165 c	156 g	180 t	1 others
ORIGIN					

Query Match	20.9%;	Score 567.4;	DB 14;	Length 699;
Best Local Similarity	88.3%;	Pred. No. 3.6e-133;		
Matches 616;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;
Qy	605	GCCTTTTCATCCAAATTCGACCTACATGAAATTCGGACATCATCTCGAGGACAATATTTTT	664	
Db	1	GCCTTCCATCCAAATCTCGACCTATTTGAAGTTGGACATCATCTCGAGGACAATATTTTT	60	
Qy	665	AAGCTGCGAGAACTCACATTTTTGAAACCGCGCTGGGAAGATGTCATATGACCCCTAAGCTG	724	
Db	61	AAGCTGCGAGACTCACATTTTGAACCGCTGGGAAGAGCCCTTATGATCCCAACTA	120	
Qy	725	CTGACAGATCCCTTTTAATAGATGGAGAAATATATCTCTGGAACCTGAGCTGATTTTATG	784	
Db	121	CTGACTCCCTCTCTTCTAATAGACGGTGAGTTGTACTCTGGAACCTGCTGGGACTTCATG	180	
Qy	785	GGGCGAGACTTGTATCTTCCGAACTCTTGGGACCAACCACCACTAAGCAGAGCAG	844	
Db	181	GGACGGGACTTCGGCTATCTTTCAGAACTCTGGGCAACCATCACCCATCAGGACGGAGCAG	240	
Qy	845	CATGATTCAGGTTGGCTCAATGATCCAAAGTTCTATTAGTCCCACTCATCTCAGAGAGT	904	
Db	241	CATGACTCCCGTGGCTCAATGATCCTTAGATTCATCAGTGGCCATCTCATCCAGAGAGT	300	
Qy	905	GACAACTCTGAAGATGACAAAGTATACCTTTTTCTTCGTGAAAATGCAATAGATGGAGAA	964	
Db	301	GACAACTCTGAAGATGACAAAGTATATTTTTCTTCCGAGAAAATGCAATAGACGGAGAA	360	
Qy	965	CACCTCTGGAAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGG	1024	
Db	361	CACCTCTGGAAAAGCCATCTATCTAGATATAGGTGAGATATGCAAGAATGACTTTGGTGG	420	
Qy	1025	CACAGAAGTCTGTTGAATAAATGACAAACATTCCTCAAAGCTCGTCTGATTTGCTCAGTG	1084	
Db	421	CACAGAAGTCTTGTGAATAAATGACAAACATTCCTTAAAGCACGCTGATTTGCTCTGTG	480	
Qy	1085	CGAGGTCCAAATGGCATTTGACATCTATTTTGTATGAACTGCGAGGATGATTCCTTAATGAAC	1144	
Db	481	CCCGGTCCCAATGGCAATGACACCAATTTTGAATGCAATGCGAGATGATTTCTTAATGAAC	540	
Qy	1145	TTTAAAGATCCTTAAATCCAGTTGTATATGGAGTGTTCAGACTTCCAGTAACATTTTC	1204	
Db	541	TCTTAAAGATCCTTAAATCCGATCCGTCTATGGAGTGTTCACACATCAAGCAACATCTTT	600	
Qy	1205	AAGGGATCAGCCGCTGTGTATGTATAGCATGAGTGTGAGAAAGGTTGTTCTTGTGTCCTA	1264	
Db	601	AAGGGATCTGCTGTGTGTCATGTACAGCATGAGTGTATGANAAGGTTGTTCTTGTGTCCTA	660	
Qy	1265	TATGCCCAAGGGATGGACCCAACTATCAATGGGTGCC	1302	
Db	661	TATGCTCACAGATGTTGCCAACTATCAGTGGGTGCC	698	

Search completed: July 31, 2003, 17:17:23
Job time : 5037 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 11:35:10 ; Search time 679 Seconds
(without alignments)
10769.923 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttcatcgatg.....aggcttttttctctaataacc 2709

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	22	AAH47049
2	2508.4	92.6	2530	19	AAX75767
3	2508.4	92.6	2530	24	ABX76512
4	2508.4	92.6	2601	16	AAQ87442
5	1415.2	52.2	1481	16	AAQ92331
6	718.6	26.5	2331	20	AAX89112
7	718.6	26.5	3871	21	AAH47049
8	718.6	26.5	3871	22	AAH47049

9	718.6	26.5	3871	22	AAF54421
10	718.6	26.5	3871	25	ACA57856
11	718.6	26.5	3871	25	ABX98326
12	718.6	26.5	3871	25	ABX98828
13	718.6	26.5	3871	25	ACA05873
14	718.6	26.5	3871	25	ABX97917
15	718.6	26.5	3871	25	ABX78701
16	718.6	26.5	3871	25	ABX75714
17	718.6	26.5	3871	25	ABX76919
18	718.6	26.5	3871	25	ABX16759
19	718.6	26.5	3880	20	AAX89113
20	647.4	23.9	6474	25	ACC50287
21	599.4	22.1	2337	22	AAF90251
22	599.4	22.1	2337	24	ABA97363
23	596.4	22.0	2898	19	AAV35367
24	594.8	22.0	4460	20	AAZ28469
25	556	20.5	5231	23	ABV23116
26	556	20.5	5231	23	ABV28955
27	554.4	20.5	2975	20	AAX03792
28	554.4	20.5	2975	21	AAC66800
29	553.6	20.4	3988	20	AAZ28470
30	552.8	20.4	5177	24	ABV94769
31	474.8	17.5	2349	22	AAH47791
32	474.8	17.5	2628	22	AAH47790
33	473.6	17.5	3568	22	AAH47792
34	473.6	17.5	4859	21	AAA93109
35	460.2	17.0	2340	22	AAF90250
36	460.2	17.0	2340	24	ABA97362
37	454.4	16.8	456	23	ABA545213
38	454.4	16.8	456	24	ABS19795
39	385.8	14.2	1444	21	AAC69627
40	374	13.8	3054	24	ABK36165
41	353	13.0	354	22	AAI29182
42	353	13.0	354	25	ABZ33368
43	335.4	12.4	496	23	ABS32213
44	335.4	12.4	496	24	ABS07290
45	225.6	8.3	561	24	ABL87727

ALIGNMENTS

RESULT 1

AAH47049

ID AAH47049 standard; DNA; 2709 BP.

XX AC AAH47049;

XX DT 29-OCT-2001 (first entry)

XX DE Semaphorin D cDNA sequence.

XX KW Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;

XX KW 24p3; proliferin; maspin; cancer; cytostatic; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200155455-A2.

XX PD 02-AUG-2001.

XX PF 31-JAN-2001; 2001WO-US03161.

XX PR 31-JAN-2000; 2000US-0179191.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI (JINS/) JIN S.

XX PI Jin S;

XX DR WPI; 2001-488799/53.

XX PT Determining if a compound modulates the drug resistance of a cell,

PT comprises determining the expression or activity level of a resistance
 XX sequence in a cell in the presence of the test compound -

PS Example 1; Fig 1A-B; 79pp; English.

XX The invention relates to a method of determining whether a test compound
 CC modulates the drug resistance of a cell that comprises determining the
 CC expression or activity level of resistance genes (e.g. semaphorin D, B94,
 CC mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
 CC the test compound, and comparing its expression or activity level in a
 CC cell without the test compound. The drug resistant sequences are useful
 CC in identifying drug resistant cells, in screening methods directed to the
 CC identification of compounds that can modulate the drug resistance of a
 CC cell type or multiple cell types. An isolated resistance protein can be
 CC used as an immunogen to generate antibodies that bind the resistance
 CC protein. Resistance nucleic acids may be inserted into vectors and used
 CC as gene therapy vectors. An anti-resistance protein antibody may be used
 CC to isolate a resistance protein, or facilitate the purification of
 CC natural resistance protein from cells and of recombinantly produced
 CC resistance protein expressed in host cells. The methods are useful for
 CC treating a subject having a disorder, such as a drug-resistance cancer,
 CC characterized by aberrant resistance sequence expression or activity by
 CC administering to the subject a resistance modulator. The present sequence
 CC represents a semaphorin cDNA sequence, whose expression was increased in
 CC drug resistant EMT6 tumours.

XX Sequence 2709 BP; 831 A; 560 C; 620 G; 698 T; 0 other;

Query Match 100.0%; Score 2709; DB 22; Length 2709;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AATCTTTTATTCGATGTTAAACAGCTTGTAGTATCATCGCCAGCTGAGGGGTGCG	60
DB	1	AATCTTTTATTCGATGTTAAACAGCTTGTAGTATCATCGCCAGCTGAGGGGTGCG	60
QY	61	ACCCAGCGCTCGGGAGTAGGTGAGCTGCGCTGTTCTCCCATTTGTCAGCCAGTCTATT	120
DB	61	ACCCAGCGCTCGGGAGTAGGTGAGCTGCGCTGTTCTCCCATTTGTCAGCCAGTCTATT	120
QY	121	CCAGATTGTTGAATCTCTCGCCGCAACATACAGGAAGGAACTAAAGCAGCAAAAGG	180
DB	121	CCAGATTGTTGAATCTCTCGCCGCAACATACAGGAAGGAACTAAAGCAGCAAAAGG	180
QY	181	GACCTACAGCGTCTGCAGCATGGCTGGTTAACTAGGATTTGCTCTTTTCTGGGGAGT	240
DB	181	GACCTACAGCGTCTGCAGCATGGCTGGTTAACTAGGATTTGCTCTTTTCTGGGGAGT	240
QY	241	ATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
DB	241	ATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
QY	301	ATCCTACAAAGAAATGTTGGAAATCCAAATGTGATCACTTTCAATGGCTTGCCCAACAG	360
DB	301	ATCCTACAAAGAAATGTTGGAAATCCAAATGTGATCACTTTCAATGGCTTGCCCAACAG	360
QY	361	CTCCAGTTATCATACCTCTCTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAA	420
DB	361	CTCCAGTTATCATACCTCTCTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAA	420
QY	421	GGATCACATATTTTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCC	480
DB	421	GGATCACATATTTTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCC	480
QY	481	AGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATG	540
DB	481	AGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATG	540
QY	541	TGCTAAATTTTCATCAAGGTAATTAAGGCATATATCAGACTCACTTGTAGCCGTGGAC	600
DB	541	TGCTAAATTTTCATCAAGGTAATTAAGGCATATATCAGACTCACTTGTAGCCGTGGAC	600
QY	601	GGGGGGCTTTTTCATCCAAATTTGCACCTCAATTGAAATTTGACATCATCTCTGAGGACAATAT	660

DB	601	GGGGGGCTTTTTCATCCAAATTTGCACCTACATGAAATTTGACATCATCTCTGAGGACAATAT	660
QY	661	TTTTAAGCTGAGAACTCACATTTTGAAGAACGGCCGTGGGAAGAGTCCATATGACCTTAA	720
DB	661	TTTTAAGCTGAGAACTCACATTTTGAAGAACGGCCGTGGGAAGAGTCCATATGACCTTAA	720
QY	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAATTAATCTCTGAACTGACAGCTGATTT	780
DB	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAATTAATCTCTGAACTGACAGCTGATTT	780
QY	781	TATGGGGCGAGACTTTTGTCTATCTTCCGAACTCTTGGGCAACCCACCACCAATCAGACAGA	840
DB	781	TATGGGGCGAGACTTTTGTCTATCTTCCGAACTCTTGGGCAACCCACCACCAATCAGACAGA	840
QY	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGA	900
DB	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGA	900
QY	901	GAGTGACAATCTCGAAGATGACAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGG	960
DB	901	GAGTGACAATCTCGAAGATGACAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGG	960
QY	961	AGAACACTCTCTGAAAAGCTACTCAGCGTAGAATAGTTCAGATATCCAAAGATGACTTTGG	1020
DB	961	AGAACACTCTCTGAAAAGCTACTCAGCGTAGAATAGTTCAGATATCCAAAGATGACTTTGG	1020
QY	1021	AGGGCAGAGAAGTCTGGTGAATTAATGCAACAACATCTCTCAAGCTCTGTGATTTGCTC	1080
DB	1021	AGGGCAGAGAAGTCTGGTGAATTAATGCAACAACATCTCTCAAGCTCTGTGATTTGCTC	1080
QY	1081	AGTCCAGCTCCAAATGGCAATTTGATCACTTTTGTGAACTGCAAGGATGATTTCTTAAT	1140
DB	1081	AGTCCAGCTCCAAATGGCAATTTGATCACTTTTGTGAACTGCAAGGATGATTTCTTAAT	1140
QY	1141	GAACTTTAAAGATCCTTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAAAT	1200
DB	1141	GAACTTTAAAGATCCTTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAAAT	1200
QY	1201	TTTCAAGGGATCAGCGGTGTGTATGATAGCATGATGATGAGAGGGTGTCTCTGG	1260
DB	1201	TTTCAAGGGATCAGCGGTGTGTATGATAGCATGATGATGAGAGGGTGTCTCTGG	1260
QY	1261	TCCATATGCCACAGGGATGACCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCC	1320
DB	1261	TCCATATGCCACAGGGATGACCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCC	1320
QY	1321	CTATCCAGGGCAGGAACTTGTCCCGCAAAACATTTTGGTGTGTTTGACTCTTCAAAAGGA	1380
DB	1321	CTATCCAGGGCAGGAACTTGTCCCGCAAAACATTTTGGTGTGTTTGACTCTTCAAAAGGA	1380
QY	1381	CCTTCTGATGATGTTTAACTTTTGCAGAAAGTATCCAGCCATGTACATCCAGTGT	1440
DB	1381	CCTTCTGATGATGTTTAACTTTTGCAGAAAGTATCCAGCCATGTACATCCAGTGT	1440
QY	1441	TCCTATGAACAATCGCCCAATAGTATCAAAACGATGTAATTAATTAATTAACAAAT	1500
DB	1441	TCCTATGAACAATCGCCCAATAGTATCAAAACGATGTAATTAATTAATTAACAAAT	1500
QY	1501	TGTGCTAGACCGAGTGGATGCAAGATGGAAGTATGATGATGATGATGATGATGATGATG	1560
DB	1501	TGTGCTAGACCGAGTGGATGCAAGATGGAAGTATGATGATGATGATGATGATGATGATG	1560
QY	1561	TGTTGGGACCGTCTTAAAGTAGTTCATTTCTTAAGGAGACTGGTATGATGATGATGATG	1620
DB	1561	TGTTGGGACCGTCTTAAAGTAGTTCATTTCTTAAGGAGACTGGTATGATGATGATGATG	1620
QY	1621	GTTTCTGCTGGAAGAAATGACAGTTTTTTCGGGAAACCGAGTCTGATTTTTCAGCAATG	1680
DB	1621	GTTTCTGCTGGAAGAAATGACAGTTTTTTCGGGAAACCGAGTCTGATTTTTCAGCAATG	1680
QY	1681	TTCCACTAAGCAGCAACACTATATATTTGGTTCAACCGCTGGGTGTCAGCTCCCTTT	1740
DB	1681	TTCCACTAAGCAGCAACACTATATATTTGGTTCAACCGCTGGGTGTCAGCTCCCTTT	1740

Db 1681 TTCCACTAAGCAGCAACAACTATATATGTTTCAAGCGGTGGGGTGGCCAGCTCCCTTT 1740
Qy 1741 ACACCGGTGTGATATTATACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTA 1800
Db 1741 ACACCGGTGTGATATTATACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTA 1800
Qy 1801 CTGTGCTTGGGATGGTTCGCAATGCTGCTGCTATTTTCCCACTGCAAGAGACGCAAG 1860
Db 1801 CTGTGCTTGGGATGGTTCGCAATGCTGCTGCTATTTTCCCACTGCAAGAGACGCAAG 1860
Qy 1861 ACACAAGATATAAGAAATGGAGACCCACTGACTCTCTCTTCCAGACTTACACCATGATAA 1920
Db 1861 ACACAAGATATAAGAAATGGAGACCCACTGACTCTCTCTTCCAGACTTACACCATGATAA 1920
Qy 1921 TCACCATGCCACAGCCCTGAAAGAGAGAAATCATCTATGTTGATAGAGATAGTACACATT 1980
Db 1921 TCACCATGCCACAGCCCTGAAAGAGAGAAATCATCTATGTTGATAGAGATAGTACACATT 1980
Qy 1981 TTTGGAATGCAGTCCGAAGTCCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAA 2040
Db 1981 TTTGGAATGCAGTCCGAAGTCCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAA 2040
Qy 2041 TGAAGAGCGAAAGAGAGATCAGAGTGGATCATATCATCAGACAGATCAAGGCT 2100
Db 2041 TGAAGAGCGAAAGAGAGATCAGAGTGGATCATATCATCAGACAGATCAAGGCT 2100
Qy 2101 TCTGCTACGTAGTCTACAAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAACA 2160
Db 2101 TCTGCTACGTAGTCTACAAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAACA 2160
Qy 2161 TGGGTTTCATAAACTCTTTTAAAGTAAACCTTGAAGTCAATGACAGAGCATTTGGA 2220
Db 2161 TGGGTTTCATAAACTCTTTTAAAGTAAACCTTGAAGTCAATGACAGAGCATTTGGA 2220
Qy 2221 AGAATCTTCTTAAGATGATGAGAGATGCTTAAGACCAAGAAATGCTCCAATAG 2280
Db 2221 AGAATCTTCTTAAGATGATGAGAGATGCTTAAGACCAAGAAATGCTCCAATAG 2280
Qy 2281 CATGACACCTAGCCAGAGGCTTGGTACAGAGACTTCATGACAGCTCATCAACACCCCAA 2340
Db 2281 CATGACACCTAGCCAGAGGCTTGGTACAGAGACTTCATGACAGCTCATCAACACCCCAA 2340
Qy 2341 TCTCAACAGATGGATGATCTGTGAAACAAGTTTGGAAAGGACCGGAAACAAACGTCG 2400
Db 2341 TCTCAACAGATGGATGATCTGTGAAACAAGTTTGGAAAGGACCGGAAACAAACGTCG 2400
Qy 2401 GCAAGGCGCAGACATACCCAGGGAACAGTAACTGGAAGCACTTACAGAAATAA 2460
Db 2401 GCAAGGCGCAGACATACCCAGGGAACAGTAACTGGAAGCACTTACAGAAATAA 2460
Qy 2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTCAGAGGGGACCCAGAGGTCTGAGCTGC 2520
Db 2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTCAGAGGGGACCCAGAGGTCTGAGCTGC 2520
Qy 2521 ATTACTCTAGAAACCTCAAAACAGTAGAAACTTGCCTAGACATTAACCTGGAACAAA 2580
Db 2521 ATTACTCTAGAAACCTCAAAACAGTAGAAACTTGCCTAGACATTAACCTGGAACAAA 2580
Qy 2581 TGCATATACATGAATTTTTCATGGCAATTTGATGATTTTCAATGTTGGAATTC 2640
Db 2581 TGCATATACATGAATTTTTCATGGCAATTTGATGATTTTCAATGTTGGAATTC 2640
Qy 2641 AGCTGAGTTCCACCAATATAAATAATCAATGAGTAACTTCTTAATAGGCTTTTTT 2700
Db 2641 AGCTGAGTTCCACCAATATAAATAATCAATGAGTAACTTCTTAATAGGCTTTTTT 2700
Qy 2701 CCTAATACC 2709
Db 2701 CCTAATACC 2709

RESULT 2
AAx75767

ID AAX75767 standard; DNA; 2530 BP.
XX AC AAX75767;
XX DT 22-JUL-1999 (first entry)
XX DE Human semaphorin III DNA.
XX KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
frameshift mutation; age-related disease; neurodegenerative disorder;
Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
Huntington's disease; multiple sclerosis; alcoholic liver disease;
diabetes mellitus type II; microtubule associated protein; tau; Big Tau;
ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPP-1;
bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
high mobility group protein-C; neuroendocrine specific protein A; ss.
XX OS Homo sapiens.
XX PN W09845322-A2.
XX PD 15-OCT-1998.
XX PF 02-APR-1998; 98WO-IB00705.
XX PR 10-APR-1997; 97US-0043163.
XX PA (UYUT-) RIJKSUNIV UTRECHT.
XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX DR WPI; 1998-609901/51.
XX PT Diagnosing disease by detecting frameshift mutations in RNA or
corresponding protein mutations - used to diagnose cancer and
neurological diseases, particularly Alzheimer's disease, and also
for treatment and prevention with specific ribozymes or wild-type
RNA
XX PS Disclosure; Figure 16; 259pp; English.
XX CC This invention describes a novel method for the diagnosis of a disease
caused by, or associated with, an RNA molecule that has a frameshift
mutation. The method is used to diagnose age-related diseases,
especially cancer and a wide range of neurodegenerative disorders (e.g.
Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's
disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus
type II and many others listed) or susceptibility to these disorders. The
method allows a definitive diagnosis of Alzheimer's disease in living
patients, at an early stage. It is based on the observation that disease
may be caused by mutations in RNA rather than DNA. The invention
describes the use of neuronal system RNA molecules, specifically
proteins including beta-amyloid precursor protein (beta-APP), the
microtubule associated proteins Tau and Big Tau, ubiquitin B,
apolipoprotein E, microtubule associated protein 2 (MAP2),
neurofilament-L, neurofilament-M, neurofilament-F, presenilin I,
presenilin II, glial fibrillary acidic protein (GFAP), the cellular
tumour antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene,
semaphorin III, HUPP-1, high mobility group protein-C (HMGP-C) and
neuroendocrine specific protein A. This sequence encodes the wild type
and mutant protein fragments represented in AAY21264-Y21348.
XX SQ Sequence 2530 BP; 786 A; 518 C; 576 G; 650 T; 0 other;
Query Match 92.6%; Score 2508.4; DB 19; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels. 0; Gaps 0;

Qy 193 CTGCAGCATGGGCTGGTTAACTAGGATTGCTGCTTTTCTGGGAGTATTACTTACAGC 252

1693 GCAACAACTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACGGGTGCA 1752
1509 GCAACAACTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACGGGTGCA 1568
1753 TATTTACGGGAAGCGTGTCTGAGTGTGGTTCCTCGCCGAGACCCCTTACTGTGCTTGGGA 1812
1569 TATTTACGGGAAGCGTGTCTGAGTGTGGTTCCTCGCCGAGACCCCTTACTGTGCTTGGGA 1628
1813 TGGTTCCTGCATGTTCTCGCTATTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1872
1629 TGGTTCCTGCATGTTCTCGCTATTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1688
1873 AAGAAATGGAGACCCACTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
1689 AAGAAATGGAGACCCACTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1748
1933 CAGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATATTTTGGAAATGCAG 1992
1749 CAGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATATTTTGGAAATGCAG 1808
1993 TCCGAAGTGCAGAGAGCGCTGCTGCTATTTGGCAATTTCCAGAGCGGAAATGAAGCGGAAA 2052
1809 TCCGAAGTGCAGAGAGCGCTGCTGCTATTTGGCAATTTCCAGAGCGGAAATGAAGCGGAAA 1868
2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 2112
1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 1928
2113 TCTAACAGAGAGGATTCAGGCAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172
1929 TCTAACAGAGAGGATTCAGGCAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1988
2173 AACTCTTCTTAAGTAACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2232
1989 AACTCTTCTTAAGTAACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2048
2233 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCAATAGATGACACCTAG 2292
2049 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCAATAGATGACACCTAG 2108
2293 CAGAGAGTCTGTGACAGAGCTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2352
2109 CAGAGAGTCTGTGACAGAGCTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2168
2353 GATGAGTCTGTGAAAGTTTGGAAAGGGACCGAAACCAACCTCGGCAAGGCCAGG 2412
2169 GATGAGTCTGTGAAAGTTTGGAAAGGGACCGAAACCAACCTCGGCAAGGCCAGG 2228
2413 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2472
2229 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2288
2473 CAGGAGGACCAAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATTAACCTCTAGA 2532
2289 CAGGAGGACCAAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATTAACCTCTAGA 2348
2533 AACCTCAAAAGTACAACTTGCCTAGACAAATACTGGAAGAAACAAATGCAATATACAT 2592
2349 AACCTCAAAAGTACAACTTGCCTAGACAAATACTGGAAGAAACAAATGCAATATACAT 2408
2593 GAACTTTTTTCATGGCAATATGTTGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2652
2409 GAACTTTTTTCATGGCAATATGTTGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2468
2653 CCAATTAATAATTAATCCATGAGTAACTTCTCTAATAGGCTTTTTTCC 2702
2469 CCAATTAATAATTAATCCATGAGTAACTTCTCTAATAGGCTTTTTTTC 2518

RESULT 5

AAQ92331.

ID AAQ92331 standard; cDNA; 1481 BP.

XX

AC AAQ92331;
XX 01-NOV-1995 (first entry)
XX Human collapsin cDNA.
XX Collapsin; antibody; therapy; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 50..1480
FT /*tag= a
XX US5416197-A.
XX 16-MAY-1995.
XX 15-OCT-1993; 93US-0136922.
XX 15-OCT-1993; 93US-0136922.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Luo Y, Raper JA;
XX WPI; 1995-193478/25.
XX P-PSDB; AAR74175.
XX New antibody to human collapsin - used to inhibit the activity of
XX collapsin, to induce neurite out-growth and to treat individuals with
XX nerve damage.
XX Disclosure; Columns 11-16; 11pp; English.
XX Human collapsin and its encoding nucleic acid may be used to
XX identify agents which modulate the ability of human collapsin to
XX collapse the growth cone of neurons. An antibody capable of
XX specifically binding at least a portion of the collapsin protein can
XX be used to purify human collapsin and to inhibit the activity of the
XX protein. It can be used to induce neurite outgrowth by neuronal
XX cells and to treat individuals suffering from nerve damage.
XX Sequence 1481 BP; 454 A; 299 C; 345 G; 383 T; 0 other;
SQ
Query Match 52.2%; Score 1415.2; DB 16; Length 1481;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 899 GAGAGTGACATCTCTGAAGATGACAAAGTATACCTTTCTTCCTGAAATGCAATAGAT 958
DB 53 GACATGACATCTCTGAAGATGACAAAGTATACCTTTCTTCCTGAAATGCAATAGAT 112
QY 959 GGAGAACACCTCTGAAAGCTACTCAGCTAGATAGTACAGATATGCAAGATGACTTT 1018
DB 113 GGAGAACACCTCTGAAAGCTACTCAGCTAGATAGTACAGATATGCAAGATGACTTT 172
QY 1019 GGAGGGACAGAGTCTGTTGAATTAATGGACACATCTCTCAAGCTCGTCTGATTTGC 1078
DB 173 GGAGGGACAGAGTCTGTTGAATTAATGGACACATCTCTCAAGCTCGTCTGATTTGC 232
QY 1079 TCAGTGGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAATGCAAGATGATTTCTTA 1138
DB 233 TCAGTGGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAATGCAAGATGATTTCTTA 292
QY 1139 ATGAACCTTTAAGATCCTTAAATAATCCAGTGTATATGAGGTGTTTACGACTTCAGTAAC 1198
DB 293 ATGAACCTTTAAGATCCTTAAATAATCCAGTGTATATGAGGTGTTTACGACTTCAGTAAC 352
QY 1199 ATTTTCAGGGATCAGCGGTGTGTATGTATATGATGATGATGATGATGATGATGATGATGAT 1258
DB 353 ATTTTCAGGGATCAGCGGTGTGTATGTATATGATGATGATGATGATGATGATGATGATGAT 412

```
QY 1259 GGTCCATATGCCACAGGAGTGAACCCAACTATCAATGGTGCCTTTATCAAGGAAGATC 1318
Db |
QY 413 GGTCCATATGCCACAGGAGTGAACCCAACTATCAATGGTGCCTTTATCAAGGAAGATC 472
Db |
QY 1319 CCTATCCACGCGCAGGACTTGTCCACGAAACATTTGGTGGTTTGGACTCTCAAG 1378
Db |
QY 473 CCTATCCACGCGCAGGACTTGTCCACGAAACATTTGGTGGTTTGGACTCTCAAG 532
Db |
QY 1379 GACCTTCCTGATGATGTTATACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTG 1438
Db |
QY 533 GACCTTCCTGATGATGTTATACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTG 592
Db |
QY 1439 TTTCTATGAACAATCGCCCAATAGTATCAAAAAGGATGTAATTAATTAACAA 1498
Db |
QY 593 TTTCTATGAACAATCGCCCAATAGTATCAAAAAGGATGTAATTAATTAACAA 652
Db |
QY 1499 ATTGTCGTAGACCGAGTGCATGACGAAGATGACAGATGATGTTATGTTATCGGAACA 1558
Db |
QY 653 ATCGTCGTAGACCGAGTGCATGACGAAGATGACAGATGATGTTATGTTATCGGAACA 712
Db |
QY 1559 GATGTTGGACCGTTCCTTAAAGTAGTTCCTTAAAGGAGACTTGGTATGATTTAGAA 1618
Db |
QY 713 GATGTTGGACCGTTCCTTAAAGTAGTTCCTTAAAGGAGACTTGGTATGATTTAGAA 772
Db |
QY 1619 GAGGTTCTGCTGGAAGAAATGACGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAG 1678
Db |
QY 773 GAGGTTCTGCTGGAAGAAATGACGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAG 832
Db |
QY 1679 CTTTCCACTAAGCAGCAACTATATATTTGTTCAACGGCTGGGGTTCGCCAGCTCCCT 1738
Db |
QY 833 CTTTCCACTAAGCAGCAACTATATATTTGTTCAACGGCTGGGGTTCGCCAGCTCCCT 892
Db |
QY 1739 TTACACCGGTGTGATATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTT 1798
Db |
QY 893 TTACACCGGTGTGATATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTT 952
Db |
QY 1799 TACTGTGCTGGGATGTTCTGCATGTTCTGCCTATTTTCCCAATGCAAGAGAGCGACA 1858
Db |
QY 953 TACTGTGCTGGGATGTTCTGCATGTTCTGCCTATTTTCCCAATGCAAGAGAGCGACA 1012
Db |
QY 1859 AGACGACAGATATAGAAATGGAGACCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCT 1918
Db |
QY 1013 AGACGACAGATATAGAAATGGAGACCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Db |
QY 1919 AATCACCATTGGCCACAGCCCTGAAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACA 1978
Db |
QY 1073 AATCACCATTGGCCACAGCCCTGAAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACA 1132
Db |
QY 1979 TTTTGGAAATGACGTCGGAAGTCGACAGAGCGCTGCTTATTTGGCAATTCAGAGGGGA 2038
Db |
QY 1133 TTTTGGAAATGACGTCGGAAGTCGACAGAGCGCTGCTTATTTGGCAATTCAGAGGGGA 1192
Db |
QY 2039 AATGAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGC 2098
Db |
QY 1193 AATGAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGC 1252
Db |
QY 2099 CTTCTGCTACGTAGTCTCAACAGAGAGGATTCAGGCAATTAATCTCTGCCATCGGTGGAA 2158
Db |
QY 1253 CTTCTGCTACGTAGTCTCAACAGAGAGGATTCAGGCAATTAATCTCTGCCATCGGTGGAA 1312
Db |
QY 2159 CATGGGTTCAACAACTCTTCTTAAAGTAAACCTTGGAAAGTCAATTTGACACAGAGATTG 2218
Db |
QY 1313 CATGGGTTCAACAACTCTTCTTAAAGTAAACCTTGGAAAGTCAATTTGACACAGAGATTG 1372
Db |
QY 2219 GAGAACTTCTTCAATAAGATGATGAGATGCTCTAAGACCCAAAGAAATGTCAT 2278
Db |
QY 1373 GAGAACTTCTTCAATAAGATGATGAGATGCTCTAAGACCCAAAGAAATGTCAT 1432
Db |
QY 2279 AGCATGACACCTAGCCAGAGGTTCTGCTACAGAGACTTTCATGCGAGCTC 2326
Db |
QY 1433 AGCATGACACCTAGCCAGAGGTTCTGCTACAGAGACTTTCATGCGAGCTC 1480
```

```
RESULT 6
AAx89112
ID AAX89112 standard; DNA; 2331 BP.
XX AC AAX89112;
XX DT 14-SEP-1999 (first entry)
XX Human brain tissue-derived polypeptide coding sequence (clone OM007).
XX Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
XX recombinant; diagnosis; treatment; ss.
XX Homo sapiens.
XX WO9933873-A1.
XX PN 08-JUL-1999.
XX PD 25-DEC-1998; 98WO-JP05952.
XX PF 26-DEC-1997; 97JP-0358811.
XX PR (ONCOY ) ONO PHARM CO LTD.
XX PA Fukushima D, Shibayama S, Tada H;
XX PI WPI; 1999-419088/35.
XX DR P-PSDB; AAY27127.
XX PT New adult human brain tissue-produced polypeptides useful for
XX diagnosis and treatment
XX Claim 4; Page 39-40; 86pp; Japanese.
XX The invention provides polypeptides (AAY27127-Y27133) produced by human
XX adult brain tissue, human bone marrow or a human umbilical cord venous
XX endothelial cell. Host cells transformed with vectors comprising the
XX nucleic acids encoding the polypeptides are used for the recombinant
XX expression of the polypeptides. The polypeptides can be used in
XX diagnosis, treatment and basic studies, with wide applications in
XX treatment depending on the activity to be aimed at. Sequences
XX AAX89112-125 represent nucleic acids encoding the polypeptides.
XX SQ Sequence 2331 BP; 723 A; 476 C; 529 G; 603 T; 0 other;
```

```
Query Match 26.5%; Score 718.6; DB 20; Length 2331;
Best Local Similarity 61.9%; Pred. No. 3.1e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGGAAGAACAAATGTCCCAAGGTGAAATTTATCTTCAAGAAATGTGGAATCCAAC 328
Db |
QY 109 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGTGCTTTCAAT 168
Db |
QY 329 AATGTGATCATTTCATGGCTTGGCCAAAGCTCCAGTTATCATCTTCTTTTGGAT 388
Db |
QY 169 AGCTGATATCTCTTTTGGTTCATCAGAGGACTGGAATTTTCAACTCTTCTTAGAT 228
Db |
QY 389 GAGGAACGGAGTGGCTGTATGTTGGAGCAAGGATCACATATTTTTCATTCGACCTGGT 448
Db |
QY 229 GAGGAAGAGGAGGCTCTCTTGGAGCCAAAGACCATCTTTCTACTCAGTCTGGT 288
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db |
QY 289 GACTTAAACAAAATTTTAAAGATTTTATTTGGCTCTGCTCAAGGAACGGTGAATTA 348
QY 506 TCGAAGTGGGTGGAAAAGACATCTGAAAGATGTGTGTAATTTTTCATCAAGTACTTAAG 565
Db |
QY 349 TGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAAATTTTCATCAGAGTACTCAG 408
QY 566 GCATATATCAGACTCATTGTACCCCTGTGGAACGGGGCTTTTCATCCCAATTTGCACC 625
Db |
QY 409 CCTATAACAAAACTCACATATATATGTGTGGAACCTGGAGCATTTTCATCCCAATATGTGG 468
```

Qy	626	TACATTTGAAATTTGGACATCATCTCTGGAGCAAAATATTTTTTAAGCTGGAGAACTCACATTTT	685
Db	469	TATATTGATCTTGGAGTCTCAAGGAGGATATATATTCAAACTAGACACACAGTAATTTG	528
Qy	686	GA AAA CGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGTGTGACAGCATCCCTTTTAATA	745
Db	529	GAGTCTGGCAGACTGAAATGCTCTTCGATCCTCAGCAGCCTTTTGCTTCAGTAATGACA	588
Qy	746	GATGGAGAAATATACCTCTGGAATCTGCAGCTGATTTATTTATGGGCGAGACTTTGCTATCTTC	805
Db	589	GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATACCTGCAITCACT	648
Qy	806	CGAACTCTTGGGC-----ACCACCACCCAATCAGGACAGCAGCATGATTCACAGG	856
Db	649	CGATCCCTTGGGCTACTCATGACCAACCACTACATCAGAATGACATTTTCAGAGCACTAC	708
Qy	857	TGGCTCAATGATCCAAAGTTTCAATAGTGGCCCACTCATCTCAGAGAGTGCAATCTCTGAA	916
Db	709	TGGCTCAATGGAGCAAAATTTATGGAACTTTCTTCATACAGACACCTACAATCCAGAT	768
Qy	917	GATGACAAAGTATATCTTTTCTCCGTGGAATATGCAATAGATCGAGNACTCTGGA AAA	976
Db	769	GATGATAAAAATATATTTCTCTTTCGTGGAATCATCTCAAGAAGCGAGTACTCTCGGATAA	828
Qy	977	GCTACTCACGCTAGAATAGTTCAGATATGCAAGAAATGCACTTTGAGGGGACACAAAGTCTG	1036
Db	829	ACCATCTCTTCGAGTTGGAAGGTTGTGAAGAAATGATGTAGAGAGCAACGACGCTG	888
Qy	1037	GTGAATAAATGGACAACTTCCTCAAAAGCTCGTCTGATTTGCTCAGTGCAGGCTCCAAAT	1096
Db	889	ATAAACAGTGGACGACTTTTCTTTAAGGCCAGACTGATTTTGCTCAATCTCTGGAAGTAT	948
Qy	1097	GGCATTCGACTCATTTTGTATGAATACTCGAGATGTATCTCTAATGAATCTTTAAAGATCT	1156
Db	949	GGGCGAGATACTTACTTTGTATGAGCTTCAAGATATTTATTTACTCCCAACAAGAGATGA	1008
Qy	1157	AAAAATCCAGTTGTATATGAGTGTTTTACGACTTCAGATAAATTTTCAAGGGATCAGCC	1216
Db	1009	AGAAATCCTGTATATGAGAGCTTTTACTTACAACCGAGCTCCATCTTCAAGAGCTCTGCT	1068
Qy	1217	GTGTGTATGTATAGCATGAGTGTGTGAGAAGGGTGTCTCTGTCATATGCCACAGG	1276
Db	1069	GTTTGTGTATAGCATGGCTGCATCAGAGCAGTTTTAAATGGTCCATATGCTCATAG	1128
Qy	1277	GATGGACCCAACTATCAATGGGTGCTTTATCAAGGAAGAGTCCCTTATCCAGCGCCAGGA	1336
Db	1129	GA AAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAATTCCTTATCCAGCGCTGGT	1188
Qy	1337	ACTTGTCCCGACAAAACATTTG---GTGGTTTGTACTCTCAAAAGAGCTTCTCGTGATGAT	1393
Db	1189	ACATGTCCAAGCAAAACCTATGACCCCACTGATTAAGTCCACCGAGATTTTTCAGATGAT	1248
Qy	1394	GTTATAACCTTTTGCAAGAAGTCAATCCAGCCATGTACAATCCAGTGTTTCTATGAACAAT	1453
Db	1249	GTCAATGATTTTATAAGCGGCCTCTGTGTATGATTAAGTCCGTATACCCAGTTGCAGGA	1308
Qy	1454	CGCCCAATATGTATCAAAAAGGATGTAATATCAATTTACAAATGTGTGTAGACCGGA	1513
Db	1309	GGACCAACGTTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGTGTGATCAT	1368
Qy	1514	GTGATGCAGAGNATGACATGATGATTTATGTTTATCGAACAGATGTTGGGACGGTT	1573
Db	1369	GTCATTCGAAGAAGATGCCAGTACGATGTAATGTTTCTTGGAAACAGACATTTGGAACGTGC	1428
Qy	1574	CTTAAAGTAGTTTCAAATCTCTAAGGACACTTGGTATGATTTAGAAGAGGTTCTGCTGGAA	1633
Db	1429	CTCAAAGTTGTACGACATTTCAAGGA AAAAGTGG---NATATGAAGAGGTAGTGTCTGGAG	1485
Qy	1634	GA AATGACAGTTTTTCGGGAACCGAGCTGCTATTTTACGAAATGGAGCTTTTCCATTAAGCAG	1693
Db	1486	GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAACTGAAATTTGCTCTGGAAGCAG	1545

```
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.

PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
PI WPI; 2000-237871/20.
XX P-PSDB; AAY99427.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
PT secreted PRO polypeptides, useful for screening of potential peptide or
PT small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2; Fig 175; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
CC transmembrane and receptor PRO proteins can be used for screening of
CC potential peptide or small molecule inhibitors of the relevant
CC receptor/ligand interactions. The polypeptides and nucleotide sequences
CC encoding them have various industrial applications, including uses as
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC PCR primers and hybridisation probes used in the isolation of the PRO
CC polypeptides from the present invention.
XX
XX Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

Query Match 26.5%; Score 718.6; DB 21; Length 3871;
Best Local Similarity 61.9%; Pred. NO. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAGACAAATGTGCCAGGCTGGAATTATCTTACAAAGAAATCTTGAATCCCAAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTCAACCTACAAAGACTTCTGCTTTCAAT 274
```

Qy	329	AATGTGATCACTTTTCAATGGCTTGGCCAA	CHAGCTCCAGTTATCATATCCTTCTTTTGGAT	388
Db	275	AGCTGTATTTCCCTTTTTGGGTTCATCAGAAG	AGCTGTATTTTCAAACTCTTCTCTTATAGAT	334
Qy	389	GAGGAAACGGAGTAGCGCTGTATGTTGGAGCA	AAAGGATCACATATTTTTCATTCGACCTGGTT	448
Db	335	GAGGAAAGAGCGAGCGCTGCTCTTGGAGCC	AAAGACCATCTTTTCTACAGCTCTGGTT	394
Qy	449	AA----TATCAAGGATTTTCAAAAGATTGTG	TGGCCAGTATCTTACACGAAAGAGATGAA	505
Db	395	GACTTAAACAAAATTTTAAAGATTTATTGG	CTGCTCAAGAGGAACGGGTGGATTA	454
Qy	506	TGCAAGTGGGTGGAAGAAGACATCCTGAA	AGAAATGTGCTAATTTTCATCAAGGTACTTAAG	565
Db	455	TGTAAATTTAGCTGGGAAGATGCCAATAC	AGAAATGTGCAAAATTTTCAACAGACTACTTCAG	514
Qy	566	GCATATAATCAGACTCACTTGTAGCCCTGT	GGACGGGCTTTTCATCCAAATTTGCACC	625
Db	515	CCCTATAACAAAATTCACATATATGTGTG	GAACCTGGAGCATTTCAATCAATATGTGGG	574
Qy	626	TACATTTGAAATTTGGACATCATCTCTG	AGGACAAATATTTTAAAGCTGGAGAACTCACATTTT	685
Db	575	TATATTGATCTTGGAGTCTAAGAGGATAT	TATATTTCAAATAGACACACATAATTTG	634
Qy	686	GAAACCGGCGTGGGAAGAGTCCATATG	ACCTTAAGCTGTGCAGAGCATCTCTTTTAAATA	745
Db	635	GAGTCTGGCAGACTGAAATGTCTTTCG	ATCTTCGATCCTCAGCAGCCTTTGCTTCAGTAAATGACA	694
Qy	746	GATGGAGAAATTATACTCTGGACCTG	AGCTGATTTTATGGGCGAGACTTTGCTATCTTC	805
Db	695	GATGAGTACCTCTACTCTGGAACAGCTT	CTGATTTCTTGGCAAGAGATACGTCAATTCAC	754
Qy	806	CGAACTCTTGGGC-----ACCACAC	CCCAATCAGGACAGCAGCATGATTTCCAGG	856
Db	755	CGATCCCTTGGCCCTACTCATGACCA	CCAATAATCAGAACTGACATTTCAGAGCACTAC	814
Qy	857	TGGCTCAATGATCCAAAGTTTCAATP	AGTCCCACTCATCTCAGAGAGTGACAACTCTGAA	916
Db	815	TGGCTCAATGGAGCAAAATTTATTTG	GAATTTCTTTCATACAGACACCTTCAATCCAGAT	874
Qy	917	GATGACAAAGTATACTTTTCTTCCGT	GAAAAATGCCAATAGATGGAGAACACTCTGGAAAA	976
Db	875	GATGATAAAATATATTTCTTCTTTT	CGTGAATCATCTCAAGAAGGACGATACCTCCGATAAA	934
Qy	977	GCTACTCACGCTAGAAATAGGTTC	AGATATGCAAGAAATGACTTTTGAGGGGCACAGAA	1036
Db	935	ACCATCTTCTCGAGTTGGAAGAGTTGT	GAAGATGATGAGAGAGCAACGACGCTG	994
Qy	1037	GTGNAATAAATGGACCAACATTCCT	CAAAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCAAAT	1096
Db	995	ATAAACAGTGGACGACTTTTCTTTA	AGCCAGACTGTATTTGCTCAATTCCTGGAAGTGAT	1054
Qy	1097	GGCATTTGACACTCATTTTGTATGA	AACTCGCAGATGTATTTCTTAATGAACATTTTAAAGATCC	1156
Db	1055	GGGCGAGATCTTACTTTTGTATG	AGCTTTCAAGATATTTTATTCTCCCAACAAGAGATGAA	1114
Qy	1157	AAAAATCCAGTTGTATATGGAGTGT	TTTACGACTTTCCAGTAAACATTTTCAAGGGATCAGCC	1216
Db	1115	AGAAATCTGTATATATGAGTCTTT	TACTTACAAACGAGCTTCAATCTTTCAAGAGCTCTGCT	1174
Qy	1217	GTGTGTATGTATAGCATGAGTGT	GTGAGAGGGTGTTCCTTGGTCCATATGCCCAAGG	1276
Db	1175	GTTTGTGTATAGCATGGCTG	CATCAGAGCAGTTTAAATGGTCCATATGCTCATAG	1234
Qy	1277	GATGGACCCAACTTATCAATGGTGC	TTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGA	1336
Db	1235	GAAAGTGCAGACCATCTGTTGGT	GTGAGTATGATGGAGAAATTCCTTATCCACGGCTCGT	1294
Qy	1337	ACTTGTCCAGCAGCAAAACATTTG--	-GTGGTTTGTACTCTCAAAGGACCTTCTCTGATGAT	1393
Db	1295	ACATGTCCAAGCAAAACCTATG	ACCCTACTGATTAAGTCCACCCGAGATTTTCAGATGAT	1354
Qy	1394	GTTATACCTTTTGCAGAAAGT	TCATCCAGCCATGTACAAATCCAGTGTTTTCTATGAACAAT	1453

Db	1355	GTCAATCAGTTTTCATAAAGCGGCACCTGTGATGTATTAAGTCGGTATACCAGTTGCAGGA	1414
Qy	1454	CGCCCAATAGTGATCAAAACGGATGTAAATTTATCAAATTTACAAAAATGTCGTAGACCGGA	1513
Db	1415	GGACCAAGCTTCAAGAGATCAATGTGGATTTACAGCTGACACAGATAGTGTGATCAT	1474
Qy	1514	GTGGATGCAGAAGATGGACAGTATGATGTATATGTTTATCGGAACAGATGTGGGACCGTT	1573
Db	1475	GTCAATGCAGAAGATGCCAGTACGATGTAAATGTTTCTTGGAACAGACATTTGGAACCTGTC	1534
Qy	1574	CTTAAAGTAGTTTCAATTCCTAAGGACACTGGTATGATTTAGNAGAGGTTCTGCTCGAA	1633
Db	1535	CTCAAAGTTGTACGACATTTCAAAGGAAAAATGG---AATATGGAAGAGGTAGTGTCTGGAG	1591
Qy	1634	GAATGACAGTTTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCCACTTAAGCAG	1693
Db	1592	GAGTTGCAGATATTCAGACACTCATCAATCATCTTGAACATGGAAATGTCTCTGAAGCAG	1651
Qy	1694	CAACAACTATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGTGAT	1753
Db	1652	CAACAAATGTACATTTGGTTCCCGAGATGGAATAGTTTCAGCTCTCTTTGCAAGATGGCAC	1711
Qy	1754	ATTTACGGGAAAGCGTGTGTGATGTTGGCTCGCCGAGACCCCTTACTGTGCTTGGGAT	1813
Db	1712	ACTTATGGGAAAGCTTCGGCAGACTGTTGTCTTGGCCAGAGACCCCTACTGTGCTCGGAT	1771
Qy	1814	GGTTCTGCATGTTCTCCGCTATTTTCCCACTGCAAGAGAGACGCAAGACACCAAGATATA	1873
Db	1772	GGAAATGCATGCTCTCGATATGCTCTTACTTCTTAAAGGAGAGCTAGACGCCAAGATGA	1831
Qy	1874	AGAAATGGAGACCCACTGACTCATCTGTTTACACTTACACCATGATATACCATGTGCCAC	1933
Db	1832	AAATATGGGACCCCAATCACCCAGTGTGGGACATCGAAGACAGCATTAGTCAATG---AA	1888
Qy	1934	AGCCCTCAAGAGAGAAATCATCTATGTTGTAGAGNATAGTAGACATTTTTTGGATGTCAGT	1993
Db	1889	ACTGCTGATGAAAGGTTGATTTTGGCAATTTGAATTTAACTCAACCTTCTTGGAAATGTATA	1948
Qy	1994	CCGAAGTCGACAGAGAGCGCTGGTCTATTGTCGAATTCAGAGGCGCAATGAAGAGCGAAAA	2053
Db	1949	CTTAATCCCAACAAAGCAACTATTAAATGTTATTCAGAGGTCAGGGATGAGCATCGA	2008
Qy	2054	GAAGAGATCAGGTGGATGATCATATCATCAGGACAGATCAAGGCCTTTGTGTACGTAGT	2113
Db	2009	GAGGAGTTGAAGCCCGATGAAGAATCATCAAAACCGGAAATATGGGTTACTGATTCGAAGT	2068
Qy	2114	CTACAAACAGAGAGATTCAGGCAATTAACCTCTGCCATGCGGTGGAAACATGGGTTTCATACAA	2173
Db	2069	TTGCAGAAGAGGATTCGTGGATGTATTACTGCAAGGCCAGGAGACACTTTTCATCCAC	2128
Qy	2174	ACTCTTCTTAAGGTAAACCTCTGGAAGTCATTTGACACAGAGCATTTTGGAAAGACTTCTTCAT	2233
Db	2129	ACCATAGTGAAGCTGACTTTTGAATGTCAATTTGAGATGAACAGATGGAAAAATACCCAGAGG	2188
Qy	2234	AAAGATGATGATG 2246	
Db	2189	GCAGCATGAGG 2201	

RESULT 8
AAS46098

AAS46098
ID AAS46098 standard: cDNA: 3871 BP.

XX AAS46098;

XX
.DT 18-DEC-2001 (first entry)XX
DE
Human DNA encoding PRO pol

xx
kw p[ro]v[er]b[is]: tumour: cancer: human: o

KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;

KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.
XX
OS Homo sapiens.
XX WO200168848-A2.
PN
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
PA (GETH) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
PI
XX WPI; 2001-602746/68.
DR P-PSDB; AAU29197.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
PS Claim 2; Fig 347; 774pp; English.
XX
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
CC primers for PRO polypeptides of the invention. The sequences of the
CC invention can be used to detect the presence of a tumour in a mammal by
CC comparing the level of expression of a PRO polypeptide in a test sample
CC of cells from the animal and a control sample of normal cells, whereby a
CC higher level of expression in the test sample indicates the presence of a
CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
CC pigs, goats and rabbits but are preferably human. The polypeptides can be

CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

Query Match 26.5%; Score 718.6; DB 22; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGACAATGTGCCAAGGCTGAAATATCTCTACAAAGAAATGTGGAAATCCCAAC 328
DB |||||
QY 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAACCTACAAAGACTTCTGCTTTCAAT 274
DB |||||
QY 329 AATGTGATCACTTTCAATGGCTTGGCCCAAGCTCCAGTTATCATACCTTCTCTTTGGAT 388
DB |||||
QY 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGGATTTTCAAACCTCTTCTCTTAGAT 334
DB |||||
QY 389 GAGGAACGGAGTAGCTGTATGTTGGACCAAGAGTACATATTTTCAATTCGACCTGGTT 448
DB |||||
QY 335 GAGGAAGAGGACGGCTGCTCTTGGGAGCCAAAGACCAATCTTTTCTACTCAGTCTGGTT 394
DB |||||
QY 449 AA---TATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGAGATGA 505
DB |||||
QY 395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCTGCTGCAAGGAACGGGTGGAATTA 454
DB |||||
QY 506 TGCAAGTGGGCTGGAAAGACATCTGAAAGAAATGTGCTAAATTTTCATCAAGTACTTAAG 565
DB |||||
QY 455 TGTAAATTTAGCTGGGAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514
DB |||||
QY 566 GCATATATCAGACTCACTTGTAGCCCTGTGGACGGGGCTTTTCATCCAAATTTGACACC 625
DB |||||
QY 515 CCTATACAAACCTCACAATATGTGTGTGGAACTGGAGCATTTTCATCCAATATGTGGG 574
DB |||||
QY 626 TACATTGAAATTTGGACATCATCTTGAGGACAATATTTTAAAGCTGGGAACTCACAATTTT 685
DB |||||
QY 575 TATATTGATCTTGGAGCTTCAAGAGGATATTTATTTCAAACTAGACACACATATTTTG 634
DB |||||
QY 686 GAAACCGCCCTGGGAAGAGTCCATATGACCCCTAAGCTGTGTCAGAGCATCCCTTTTAATA 745
DB |||||
QY 635 GAGTCTGGCAGACTGAAATGTCCTTTCGATCCTCAGCAGCCTTTTTCCTCAGTAATGACA 694
DB |||||
QY 746 GATGAGAAATTAATCTCTGGAACCTGCAGCTGATTTTATGGGGCGAGACTTTGCTATCTTC 805
DB |||||
QY 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGATCTGCATTTCACT 754
DB |||||
QY 806 CGAATCTTGGGC-----ACCACCCCAATCAGGACAGAGCAGCATGATTTCCAGG 856
DB |||||
QY 755 CGATCCCTTGGGCCCTCTCATGACCAACCTACATCAGAACTGACATTTTCAGAGCACTTAC 814
DB |||||
QY 857 TGGCTCAATGATCCAAAGTTTCAATAGTGCCCACTCATCTCAGAGAGTGACAATCCTGAA 916
DB |||||
QY 815 TGGCTCAATGAGCAAAATTTTGGAACTTCTTTCATACCAGACACCTCAATCCAGAT 874
DB |||||
QY 917 GATGACAAAGTATATCTTTTCTTCCTCGTGAATGCAATAGATGGAGAACACTCTGGAAAA 976
DB |||||
QY 875 GATGATAAAATATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 934
DB |||||
QY 977 GCTACTCAGCTAGAAATAGTGCATATGCAAGATGACTTTTGGAGGGCAGAGATCTG 1036
DB |||||
QY 935 ACCATCTTTCTCGAGTTGGAAGAGTTTGTGAAGATGATGTAGGAGGACAAACGGACCTG 994
DB |||||
QY 1037 GTGAATAAATGGACAAATCTCCTCAAAGCTCGTCTGATTTTCTCAGTCCAGGTCCTCAAT 1096
DB |||||
QY 995 ATAAACAGTGGACGACTTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGGAGTAT 1054
DB |||||
QY 1097 GGCATTGACACTCAATTTTGATGAATCGCAGGATGATTTCTTCAATGAATCTTTAAAGATCCT 1156
DB |||||

Db 1055 GGGGCGAGATACCTACTTTGATGAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
QY 1157 AAAAAATCCAGTTGTATATATGAGTGTATTCAGCTTCAGTAAATATTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCCTGTAGTATATGAGTCTTTACTAACCAAGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGTATGTATAGATGATGTGAGAGGGTGTCTTGGTCCATATGCTCCACAGG 1276
Db 1175 GTTGTGTGTATAGCTGCTGATCAGAGCAGTGTATTAATGGTCCATATGCTATAG 1234
QY 1277 GATGACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCAAGGCGGAGGA 1336
Db 1235 GAAAGTGCAGACCATCGTTGGGTGAGTATGATGGGAGAAATTCCTTATCCAGGCGTGT 1294
QY 1337 ACTTGTCCAGCAAAACATTTG---GTGGTTTTGACTCTCAAGAGGACCTTCTGATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAAGTCAATCCAGCATGTCAATCCAGTGTCTTCTATGAACAAT 1453
Db 1355 GTCATCAGTTTATAAAGCGGCACTCTGTGATGTATAGTCCGTATACCCAGTTGCAGGA 1414
QY 1454 CGCCCAATAGTCAATCAAAACGAGTAAATATCAATTTACAAATTTGCTGTAGACCGGA 1513
Db 1415 GGACCAACCTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGGTGGATCAT 1474
QY 1514 GTGGATGCAAGATGGAAGTATGATGTATGTTTATCGGAACAGATGTTGGGACCGTT 1573
Db 1475 GTCATTTGCAAGAGATGGCCAGTACGATGATGTTTCTTGGAAACAGACATTTGGAATGTC 1534
QY 1574 CTTAAAGTATGTTTCAATCTTAAGGAGACTGTGTATGATTTAGAGAGGTTCTGCTGGAA 1633
Db 1535 CTCAAAGTTGTGAGCATTTCAAGGAAGAGTGG---AATATGGAAGAGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAACGACTGCTATTTACAGCAATGAGAGCTTTCCACTAAGCAG 1693
Db 1592 GAGTTGCAGATATCAAGCACTCATCAATCATCTTGAACATGGAATTTGCTCTGAAGCAG 1651
QY 1694 CAACACTATATATTTGTTCAACGGCTGGGTGCGGCTCCCTTTACACCGGTGAT 1753
Db 1652 CAACAAATGTAATTTGTTCCGAGATGATGATTTAGTTCAGCTCTCTTGCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTCTGAGTGTGCTCGCTCGCCGAGAGCCCTTACTGTCTTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTGCAGAGAGCCCTACTGTCTGCTGGAT 1771
QY 1814 GGTCTGCAATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGACGACAAAGATATA 1873
Db 1772 GGAATATGCAATGCTCTGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
QY 1874 AGAATGGAGACCCACTGACTCTTTCAGACTTACACCATGATTAATCACCATGCCAC 1933
Db 1832 AAATATGGCGACCAATCAACCGAGTCTGGGACATCGAAGACAGCATTAAGTATG---AA 1888
QY 1934 AGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACATTTTGAAGTCACT 1993
Db 1889 ACTGCTGATGAAGAGTGAATTTTGGCATTTGAATTTAACTCAACCTTCTGGAATGATA 1948
QY 1994 CCGAAGTCCGAGAGCGCTGTCTATTGGCAATTCAGAGCGGCAAAATGAAGAGCGAATA 2053
Db 1949 CCTAAATCCCAAGCAACTATTAAATGGTATATCCAGAGGTCAGGGGATGAGCATCGA 2008
QY 2054 GAAGATCAGATGAGTGAATCATATCATCAGACAGATCAAGGCTTCTGCTAGTGT 2113
Db 2009 GAGGAGTTCAAGCCCGATGAAGAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT 2068
QY 2114 CTACAAAGAGGATTCAGCAATTTACCTCTGCGATGCGGTGGAAACATGGGTTTCATCAA 2173
Db 2069 TTGCAAGAGAGGATTTCTGGATGTATTACTGCAAGCCCGAGGACACATTTTCATCCAC 2128
QY 2174 ACTCTTTAAAGGTAAACCTCGGAAGTCAATGACAGAGCATTTTGAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGATGTTGATTTGATGAAATGAACAGATGGAATAATCCAGAGG 2188

QY 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 9
AAF54421
ID AAF54421 standard; DNA; 3871 BP.
XX
AC AAF54421;
XX
DT 02-APR-2001 (first entry)
XX
DE DNA encoding protein of the invention #86.
XX
KW Secreted; transmembrane; gene therapy; ss.
XX
OS Unidentified.
XX
PN WO200078961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US04342.
XX
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
XX
(GETH) GENENTECH INC.

XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski FJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;
PI Watanabe CK, Williams PM, Wood WI;
XX WPI; 2001-071395/08.

XX Secreted and transmembrane proteins and nucleic acids designated PRO,
-PT useful as hybridization probes, in chromosome and gene mapping and gene
therapy -

XX Claim 2; Fig 171; 787pp; English.

XX The present invention relates to secreted and transmembrane proteins.
CC These proteins and the DNA encoding them may be used as hybridization
CC probes, in chromosome and gene mapping and in the generation of
CC anti-sense RNA and DNA. They may also be used to generate either
CC transgenic animals or knockout animals which are in turn useful for
CC development and screening of therapeutically useful reagents.
CC The nucleic acids may also be used in gene therapy.

XX Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

Query Match 26.5%; Score 718.6; DB 22; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGACAATGTGCCAAGGCTGAAATTTATCTACAAAGAAATGTTGGAATCCAAAC 328
Db 215 ACTTTGAAGCAAAATATTTCCAAAGACTCAAGCTAACCTCAAGAGCTTGCTGCTTCAAAAT 274
QY 329 AATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTCATACCTTCTCTTTGGAT 388
Db 275 AGCTGATATCCCTTTTGGGTTTCATCAGAGGAGCTGGGATTTTCAAAACTCTTCTCTAGAT 334

PN US2003036143-A1.
XX 20-FEB-2003.
XX PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32878.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-US9263P.
PR 18-SEP-1997; 97US-US9266P.
PR 17-OCT-1997; 97US-US92250P.
PR 21-OCT-1997; 97US-US63486P.
PR 24-OCT-1997; 97US-US63120P.
PR 24-OCT-1997; 97US-US63121P.
PR 28-OCT-1997; 97US-US63540P.
PR 28-OCT-1997; 97US-US63541P.
PR 28-OCT-1997; 97US-US63544P.
PR 28-OCT-1997; 97US-US63564P.
PR 29-OCT-1997; 97US-US63734P.
PR 31-OCT-1997; 97US-US63870P.
PR 31-OCT-1997; 97US-US64103P.
PR 13-NOV-1997; 97US-US65311P.
PR 21-NOV-1997; 97US-US66120P.
PR 24-NOV-1997; 97US-US66466P.
PR 24-NOV-1997; 97US-US66772P.
PR 11-DEC-1997; 97US-US69335P.
PR 12-DEC-1997; 97US-US69425P.
PR 17-DEC-1997; 97US-US69870P.
PR 18-DEC-1997; 97US-US68017P.
PR 10-MAR-1998; 98US-US77450P.
PR 11-MAR-1998; 98US-US77632P.
PR 11-MAR-1998; 98US-US77649P.
PR 20-MAR-1998; 98US-US78886P.
PR 20-MAR-1998; 98US-US78939P.
PR 27-MAR-1998; 98US-US79664P.
PR 27-MAR-1998; 98US-US79786P.
PR 31-MAR-1998; 98US-US80107P.
PR 31-MAR-1998; 98US-US80194P.
PR 01-APR-1998; 98US-US80327P.
PR 01-APR-1998; 98US-US80333P.
PR 08-APR-1998; 98US-US81049P.
PR 08-APR-1998; 98US-US81070P.
PR 09-APR-1998; 98US-US81195P.
PR 15-APR-1998; 98US-US81838P.
PR 21-APR-1998; 98US-US82568P.
PR 21-APR-1998; 98US-US82569P.
PR 22-APR-1998; 98US-US82704P.
PR 22-APR-1998; 98US-US82797P.
PR 28-APR-1998; 98US-US83222P.
PR 28-APR-1998; 98US-US83495P.
PR 29-APR-1998; 98US-US83496P.
PR 29-APR-1998; 98US-US83499P.
PR 29-APR-1998; 98US-US83559P.
PR 05-MAY-1998; 98US-US84366P.
PR 06-MAY-1998; 98US-US84414P.
PR 07-MAY-1998; 98US-US84639P.
PR 07-MAY-1998; 98US-US84640P.
PR 07-MAY-1998; 98US-US84643P.
PR 15-MAY-1998; 98US-US85579P.
PR 15-MAY-1998; 98US-US85580P.
PR 15-MAY-1998; 98US-US85582P.
PR 15-MAY-1998; 98US-US85700P.
PR 18-MAY-1998; 98US-US86023P.
PR 22-MAY-1998; 98US-US86392P.
PR 22-MAY-1998; 98US-US86486P.
PR 28-MAY-1998; 98US-US87098P.
PR 28-MAY-1998; 98US-US87208P.
PR 02-JUN-1998; 98US-US87609P.
PR 02-JUN-1998; 98US-US87759P.
PR 03-JUN-1998; 98US-US87827P.
PR 04-JUN-1998; 98US-US88025P.
PR 04-JUN-1998; 98US-US88028P.
PR 04-JUN-1998; 98US-US88029P.
PR 04-JUN-1998; 98US-US88033P.
PR 04-JUN-1998; 98US-US88326P.
PR 05-JUN-1998; 98US-US88167P.
PR 05-JUN-1998; 98US-US88202P.
PR 05-JUN-1998; 98US-US88212P.
PR 05-JUN-1998; 98US-US88217P.
PR 09-JUN-1998; 98US-US88555P.
PR 10-JUN-1998; 98US-US88722P.
PR 10-JUN-1998; 98US-US88738P.
PR 10-JUN-1998; 98US-US88740P.
PR 10-JUN-1998; 98US-US88811P.
PR 10-JUN-1998; 98US-US88824P.
PR 10-JUN-1998; 98US-US88825P.
PR 10-JUN-1998; 98US-US88826P.
PR 11-JUN-1998; 98US-US88861P.
PR 11-JUN-1998; 98US-US88863P.
PR 11-JUN-1998; 98US-US88876P.
PR 12-JUN-1998; 98US-US89090P.
PR 12-JUN-1998; 98US-US89105P.
PR 16-JUN-1998; 98US-US89512P.
PR 16-JUN-1998; 98US-US89514P.
PR 17-JUN-1998; 98US-US89538P.
PR 17-JUN-1998; 98US-US89598P.
PR 17-JUN-1998; 98US-US89653P.
PR 18-JUN-1998; 98US-US89908P.
PR 19-JUN-1998; 98US-US89952P.
PR 22-JUN-1998; 98US-US90246P.
PR 22-JUN-1998; 98US-US90252P.
PR 22-JUN-1998; 98US-US90254P.
PR 24-JUN-1998; 98US-US90429P.
PR 24-JUN-1998; 98US-US90435P.
PR 24-JUN-1998; 98US-US90444P.
PR 24-JUN-1998; 98US-US90461P.
PR 24-JUN-1998; 98US-US90535P.
PR 24-JUN-1998; 98US-US90540P.
PR 25-JUN-1998; 98US-US90676P.
PR 25-JUN-1998; 98US-US90678P.
PR 25-JUN-1998; 98US-US90688P.
PR 25-JUN-1998; 98US-US90690P.
PR 25-JUN-1998; 98US-US90694P.

PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098116P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.
PR 10-SEP-1998; 98US-099812P.
Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTCAAGAAAGAAATGTGTGAAATCCCAAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTAACCTAACAAAGACTTGTCTGCTTTCAAT 274
QY 329 AATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCTCTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACTCTTCTCTTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTTCGACCTGGTT 448
DB 335 GAGGAAGAGGAGGCTGCTCTTGGAGCCAAAGACCACTCTTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAAAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAGAGATTTATTTGGCTGTCTGCAAGGAACGGGTGAATTA 454
QY 506 TCGAAGTGGGCTGGAAGAGACATCTCTGAAAGATGTGCTAAATTTTCATCAAGGTACTTTAAG 565
DB 455 TGTAAATAGCTGGGAAGAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATATCATGACTCAGTCTGTAGCTGTGGAAGGGGGCTTTTCATCCAAATTTTGACC 625
DB 515 CCTATACAAAACCTCAGATATGTGTGGAACTGGAGCAATTCATCCAAATATGTGG 574
QY 626 TACATTTGAATTTGGACATCATCTCAGGACCAATATTTTAAAGCTGGGAACCTCACATTTT 685
DB 575 TATATTGATCTGGAGTCTCAAGGAGGATTTATTTTCAAACTTAGACACACATAAATTTG 634

QY 686 GAAAAAGCGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTTCGATCTCTCAGCAGCCTTTTGTCTTTCAGTAATGACA 694
QY 746 GATGAGAAATATATCTCTGGAAGTCTGAGTGTATTTTATGGGCGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGTATTTCTTGGCAAGATACATGCAATCACT 754
QY 806 CGAACTCTTTGGC-----ACCACCAACCAATCAGGACAGAGCAGCATGATCCAGG 856
DB 755 CGATCCCTTTGGCCTACTCATGACCAACCATACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAACATCTCGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTTATGGAATCTTCTTCTCATACCAAGACACCTTACAATCCAGAT 874
QY 917 GATGACAAAGTATATCTTTCTTCGTGAAATGCAATAGATGGAGAAACACTCTCGAA 976
DB 875 GATGATAAAATATATTTCTTCTTCGTGAATCATCTCAAGAAAGGAGTACCTCCGATAA 934
QY 977 GCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTGGAGGGGACAGAAAGTCTG 1036
DB 935 ACCATCTTTCTCGAGTTGGAGAGATTTGTAAAGATGATGAGGAGGACACGCGACCTG 994
QY 1037 GTGAATAAATGGACAAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAAT 1096
DB 995 ATAAACAAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATCTCTGGAAGTGT 1054
QY 1097 GGCATTGACACTCATTTTGTGAATGAACTGCAGAGATGTATTCCTTAATGAACCTTTAAGATCT 1156
DB 1055 GGGCAGAGATCTTACTTTGTAGTGGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
QY 1157 AAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCC 1216
DB 1115 AGAAATCTGTAGTATATGAGTCTTTACTAACACAGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGTATGTATAGCATGAGTGTGTGAGAGGGTGTCTTGTGTCATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTTATATGTTCCATATGCTCATAG 1234
QY 1277 GATGACCCCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCGCAGGA 1336
DB 1235 GAAAGTCGAGACCACTCGTTGGGTGAGTATGATGGGAGAAATTCCTTATCCAGGCTGCT 1294
QY 1337 ACTGTCTCCAGCAAAAATTTG---GTGGTTTTCAGCTCTCAAAAGGACTTCTCTGATGAT 1393
DB 1295 ACATGTCCAGCAAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTCTTCTATGAACAA 1453
DB 1355 GTCATCAGTTTCAATAAGCGGCACTCTGTGATGTATAAGTCCGTATATCCAGTTGCGAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTGTAGACCGA 1513
DB 1415 GGACCAACGTTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGTGGATCAT 1474
QY 1514 GTGGATCAGAGAGATGACAGTATGTATGTATTTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTCAATCCAGAGATGCCAGTACGATGTAATTTCTTGGAAACAGACATTTGGAACTGTC 1534
QY 1574 CTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCTGCTGAA 1633
DB 1535 CTCAAAGTTGTGACATTTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTTTCCGGAACCGACTGTCTATTTTCAGCAATGGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTCAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTGTCTCTGAAGCAG 1651
QY 1694 CAACCACTATATTTGTTTCAACGGCTGGGGTCCCGAGCTCCCTTTTACACCGGTGTGAT 1753
DB 1652 CAACAATTTGATGTTTCCCGAGATGGATAGTTAGTTCAGCTCTCTCTTGCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGGTGTGCTGAGTGTGTGCTCGCCGAGACCCCTTACTGTCTGGAT 1813

```
Db 1712 ACTATGGGAAGCTGGCGAGACTGTGCTTGGCAGAGACCCCTACTGTGCTGGGAT 1771
Qy 1814 GGTCTGCAATGTTCTCGCTATTTTCCACTGCAAGAGACGCAAGATATA 1873
Db 1772 GGAATATGCAATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGGAGACCCACTGACTCTCTCTGCTTACACTTACACCATGATATCACCATGCGCCAC 1933
Db 1832 AATATGGGACCCCAATCACCAGTCTGGACATCGAAGACGATTAAGTATG---AA 1888
Qy 1934 AGCCCTGAAGAGAGAATCATCTATGTTGTAGAGATAGTAGCACATTTTGGAAATGCACT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCAATGAACTCAACCTTCTGGAATGATA 1948
Qy 1994 CGAAGTCGAGAGAGCGGTGCTATTTGGCAATTCAGAGCGGGAATGAAGAGCGAAA 2053
Db 1949 CCTAAATCCCAACAAGCAACTATTAAATGATATATCCAGAGGTGAGGGATGAGCATCGA 2008
Qy 2054 GAAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCCCTTCTGCTACGTAGT 2113
Db 2009 GAGAGTTGAAGCCCGATGAAGAATCATCAAAACGAATATGGGCTACTGATTCGAAT 2068
Qy 2114 CTACACAGAGAAGATTACGCAATTAACCTCTGCTGCGTGGAAACATGGGTTTCATACAA 2173
Db 2069 TTGCACAAGAGGATCTTGGATGTATTACTGCAAAGCCCGAGGACACACTTTCATCCAC 2128
Qy 2174 ACTCTCTTAAGTAACCTGGAAGTCATGACAGAGAGATTTGGAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGAATGTCAATGAGATGAACAGATGGAATAATACCCAGAGG 2188
Qy 2234 AAGATGATCATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 11
ABX98326
ID ABX98326 standard; cdna; 3871 BP.
XX
AC ABX98326;
XX
DT 19-MAY-2003 (first entry)
DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 347.
XX
KW Human; ss; gene; PRO; secreted protein; transmembrane protein;
KW cytosolic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188767.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US11705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
```

[illegible]

Db 935 ACCATCTTCTCGAGTGGAGAGTTTGTAAAGATGATGTAGGAGCAACAGCCGCTG 994
Qy 1037 GTCAATAAATGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCAAT 1096
Db 995 ATAAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAATGAT 1054
Qy 1097 GGCATTGACACTCATTTTGTATGAATCGCAGGATGATTCCTAATGAACCTTTAAAGATCCT 1156
Db 1055 GGGGAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCAACAGAGATGA 1114
Qy 1157 AAAATCCAGTTGATATGAGAGTGTATACGACTTCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATATGAGTCTTTACTACAAACAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATATAGCATGATGTGAGAGGGTGTCTTGGTCCATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGCTGACATCAGACGAGTTTAAATGGTCCATATCTCATAAG 1234
Qy 1277 GATGACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGA 1336
Db 1235 GAAAGTGCAGACCATCGTTGGGTGAGTATGATGGAGATTCCTTATCCAGGCTGT 1294
Qy 1337 ACTTGTCCAGCAAAACATTTG---GTGGTTTGACTCTACAAAGAACCTTCTCTGATGAT 1393
Db 1295 ACATGTCCAAAGCAAAACCTATGACCCACTGATTAAGTCCACCCAGATTTTCCAGATGAT 1354
Qy 1394 GTTATACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATGTTCATAAAGCGGCACTCTGTGATGATTAAGTCCGTATACCCCATGTCAGGA 1414
Qy 1454 CGCCCAATAGTATCAAAACGAGATGAAATTTATCAATTTACACAAATTCGTGAGACGA 1513
Db 1415 GGACCAACGTTCAAGAGATCAATGTGGATTACAGACTGACACAGATAGTGTGATCAT 1474
Qy 1514 GTGGATGCAAGAGATGGAAGATGATGATGTTTATCGGAACAGATTTGGGACCGPT 1573
Db 1475 GTCAATGCAAGAGATGGCCAGTACGATGTAATGTTCTTGGAAACAGACATTCGAATGTC 1534
Qy 1574 CTTAAAGTATGTTCAATCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA 1633
Db 1535 CTCAAAGTGTGACAGATTTCAAGGAAAGTGG---AATAGGAAGAGTAGTGTGGAG 1591
Qy 1634 GAAATGACAGTGTTCGGGAACCGACTGTATTTACAGCAATGAGCTTTCCACTAAGCAG 1693
Db 1592 GAGTGCAGATTTCAAGCACTCATCAATCACTTGAACATGGAATGTCTCTGAGCAG 1651
Qy 1694 CAACAACTATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACCCGGTGTGAT 1753
Db 1652 CAACAACTTACATTTGGTTCCGAGATGGATGATTTAGTTCAGCTCTCTTGCACAGATGGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTCTGATGTTGCTCGCCGAGACCTTTACTGTCTGCTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCGAGACTGTTGCTTGGCAGAGACCCCTTACTGTGCTGGGAT 1771
Qy 1814 GGTTCGTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGACAGACCAAGATATA 1873
Db 1772 GGAATGCACTGCTCTCGATATGCTCTACTCTTAAAGGAGAGCTAGACGCCAAGATGA 1831
Qy 1874 AGAATGGAAGCCCACTGACTGCTGTTGAGATTTACACATGATTAATCACCATGGCCAC 1933
Db 1832 AAATATGGGACCCCAATCACCAGTCTGGGACATCGAAGACAGCATAGTATG---AA 1888
Qy 1934 AGCCCTGAGAGAGATCATCTATGTTGTAGAGATAGTAGCACATTTTGAATGCAGT 1993
Db 1889 ACTGCTGATGAAGAGTGAATTTTGGCATTTGAATTTAACTCAACCTTCTGGAATGTATA 1948
Qy 1994 CGCAAGTCCAGAGCGCTGTGTTCTATTGGCAATTCAGAGGCGAAATGAAGAGCGAAA 2053
Db 1949 CCTAAATCCCAACAGCACTATTAATGGTATATCCAGAGTCCAGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGATGATCATATCATCAGACAGATCAAGGCGCTTCTGCTTACGTAGT 2113

Db 2009 GAGGAGTTGAAGCCCGATGAAGAATCATCAAAACGGAATATGGCTACTGATTCGAAGT 2068
Qy 2114 CTACAACAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAACATCGGTTTACATAA 2173
Db 2059 TTGCAAGAAGAGATTCGGGATGTTACTTGAAGCCAGGACACACTTTTCAATCCAC 2128
Qy 2174 ACTCTTCTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCATTTTGGAGAAGACTTTCTTCA 2233
Db 2129 ACCATAGTGAAGTGCATTTGAAATGTCATTTGAGATGAACAGATGGAATAATACCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 12
ABX98828
ID ABX98828 standard; cDNA; 3871 BP.
XX AC ABX98828;
XX 20-MAY-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO1491 cDNA.
KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour;
KW gene; ss.
XX OS Homo sapiens.
XX PN US2003036157-A1.
XX PD 20-FEB-2003.
XX PF 02-JUL-2002; 2002US-0188769.
XX PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.

PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063240P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065111P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 15-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081338P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084539P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087609P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.

PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089553P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090698P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 01-JUL-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.

Query Match

26.5%; Score 718.6; DB 25; Length 3871;

Best Local Similarity 61.9%; Pred. No. 4.2e-206; Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;	
QY	269 AATGGAAGACAAATGTCGAGGCTGAATATCTACAAAGAAATGTTGGAATCCAC 328
Db	215 ACTTTGAAGCAAAATATCCAGACTCAAGCTAAACCTACAAAGACTTGTCTTTCAAAT 274
QY	329 AATGTGATCACATTTCAATGGCTTGGCCAAACAGCTTATCATACCTTCTCTTTGGAT 388
Db	275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGNCTGGATTTTCAAACTCTTCTTTAGAT 334
QY	389 GAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCAATATTTTCATTCGACCTGGTT 448
Db	335 GAGGAAGAGGCGAGGCTGCTCTTGGAGCCAAAGACCACATCTTTCTACTCAGCTGGTT 394
QY	449 AA---TATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db	395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCTGCTGCAAGGAACGGGTGGAATTA 454
QY	506 TGCAAGTGGGCTGGAAGACATCTCTGAAGAAATGCTAAATTTTCATCAAGGTACTTAAG 565
Db	455 TGTAATAGCTGGGAAGATGCGCAATACAGAAATGTCAAATTTTCATCAGAGTACTTCAG 514
QY	566 GCATATAATCAGACTCACCTTGTACGCTGTGGAACGGGGCTTTTTCATCCAATTTGCACC 625
Db	515 CCTATAACAAAACTCACATATATGTGTGGAACCTGGAGCATTTTCATCCAATATGTGGG 574
QY	626 TACATTTGAAATTTGGACATCATCTCTGAGGACAATATTTTAAAGTGGAGAACTCACATTT 685
Db	575 TATATTTGATCTTTGGAGTCTACAGGAGGATATATATTTCAAACTAGACACATAATTTG 634
QY	686 GAAACGGGCTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCTCTTTTAATA 745
Db	635 GAGTCTGGGAGACTGAAATGTCTTTTCGATCTCAGCAGCCTTTTGTCTCAGTAATGACA 694
QY	746 GATGGAATATATCTCTGGAATCGAGCTGATTTTATGGGCGAGACTTTTCTATCTTC 805
Db	695 GATGAGTACTCTCTCTCTGGAACAGCTTCTGATTTCTTTGGCAAGATACGCAATCACT 754
QY	806 CGNACTCTTGGGCT 856
Db	755 CGATCCCTTTGGGCT 814
QY	857 TGCTCAATGATCCAAAGTTTCTATGTCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 916
Db	815 TGCTCAATGGAACAAAATTTATGGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 874
QY	917 GATGACAAAGTATATCTTTCT 976
Db	875 GATGATAAAATATATTTCT 934
QY	977 GCTACTCAGCTAGATAGTCTAGATATGCAAGATGACTTTGGAGGCGACAGAGTCTG 1036
Db	935 ACCATCTTTCTCGAGTTGGAAGAGTTTGTAAAGATGATGTAGGAGGACAACGCGACCTG 994
QY	1037 GTGAATAAATGGAACATTTCTCTCAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1096
Db	995 ATAAACAGTGGACGACTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTGGAAGTAT 1054
QY	1097 GGCATTGACACTCATTTTGTGAACCTGCGAGTATCTCTTAATGAACCTTTAAAGATCCT 1156
Db	1055 GGGCAGATATCTTTCTGATGAGCTTCAAGATATTTATTTACTCCCAACAGAGATGAA 1114
QY	1157 AAAATCCAGTTGTATAGGAGTGTATTCAGCTTCCAGTAAATTTTCAAGGATCAGCC 1216
Db	1115 AGAAATCCCTGTATATAGGAGTCTTTTACTAACCAGCTCCATCTTCAAGGCTCTGCT 1174
QY	1217 GTGTGTATGTATAGCATGATGTGAGAGGGTGTCTCTTGGTCCATATGCCACAGG 1276
Db	1175 GTTTGTGTATAGCATGCTGACATCAGAGCGATTTTAAATGGTCCATATGCTCATAG 1234
QY	1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTCTATCCAGCGCAGGA 1336
Db	1235 GAAAGTGCAGACCATCTGTTGGGTGCAGTATGATGGGAGAAATTCCTTATCCACGGCTGT 1294
QY	1337 ACTTGCTCCAGCAAAACATTTG---GTGTTTTGACTCTACAAAGAGACTTCTCTGATGAT 1393
Db	1295 ACATGCTCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY	1394 GTTATAACCTTTTGCAGAAGTCAATCCAGCATGTACATCCAGTGTTCCTATGAACAT 1453
Db	1355 GTCATCAGTTTCATAAAGCGGCACTCTCTGTATGATTAAGTCCGTATACCAGTTGCAGGA 1414
QY	1454 CGCCCAATAGTATCAAAACCGATGTAATTTATCAATTTACAAATTTGTCTGTAGACGA 1513
Db	1415 GGACCAACGTTTCAAGAGAAATCAATGTGTGATTAACAGACTGACACAGATGTGTGATCAT 1474
QY	1514 GTGGATGCAGAAGATGGAACAGTATGATGTTTATTCGGAAACAGATGTTTGGGACCGTT 1573
Db	1475 GTCATTTGAGAAGATGGCCAGTACGATGATGTTTCTTTGGAACAGACATTTGGAATCTGTC 1534
QY	1574 CTTAAAGTAGTTTCAATTTCTTAAGGAGACTTTGGTATGATTTTAGAAGAGGTTTCTGCTGAA 1633
Db	1535 CTCAAAGTTGTGCAATTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY	1634 GAAATGACAGTTTTCGGGAACCGGACTGCTATTTTTCAGCAATGAGAGCTTTTCCATGAGCAG 1693
Db	1592 GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTTGAACATGGAATTTGTCTCTGAAGCAG 1651
QY	1694 CAACAACTATATATTTGTTTCAACGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db	1652 CAACAAATTTGATTTGTTTCCGAGATGGAATTTAGTTTCTCTCTCTCTCTCTCTCTCTCTCT 1711
QY	1754 ATTTACGGGAAGGCTGTCTGAGTGTTCCTCGCCGAGAGCCCTTACTGTCTTTGGGAT 1813
Db	1712 ACTTATGGGAAGCTTTGCGAGACTGTTGTTCTTGCAGAGAGCCCTCTACTGTCTCTGGAT 1771
QY	1814 GGTCTGCAATGTTCTCGCTATTTTCCCACTGGAAGAGACCGCAAGAGCAGCAATATA 1873
Db	1772 GGAATGCACTGCTCTCGATATGCTCTACTTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
QY	1874 AGAATGGAAGACCCAGCTGACTGCTCTGCTTACACCTTACACCATGATTAATCACCATGCCAC 1933
Db	1832 AAATATGGCGACCCCAATCACCAGTGTGGGACATTCGAAGACAGCATTTAGTCTAG---AA 1888
QY	1934 AGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACATTTTGTGAATGCAAT 1993
Db	1889 ACTGCTGATGAAGAGTGTATTTTGGCAATTTAACTCAACCTTTCTGGAATGTATA 1948
QY	1994 CCGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 2053
Db	1949 CTTAAATCCCAACAGCAACTATTTAAATGGTATATCCAGAGGTCAAGGAGTGAAGCATCGA 2008
QY	2054 GAAGAGATCAGATGCGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGT 2113
Db	2009 GAGGAGTTGAAGCCCGATGAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT 2068
QY	2114 CTACAACAGAAAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTCATCAA 2173
Db	2069 TTGCAAGAGAGGATTTCTGGGATGTATTTACTGCAAGGCCAGGAGACACATTTTCATCCAC 2128
QY	2174 ACTCTTTTAAAGGTAACCTCGAAGTCAATGACACAGAGCATTTTGGAGAACTTCTTCAT 2233
Db	2129 ACCATAGTGAAGCTGACTTTGAAATGCTATTGAGAATGAACAGATGGAATAATACCCAGAGG 2188
QY	2234 AAAGATGATGATG 2246
Db	2189 GCAGAGCATGAGG 2201
RESULT 13 ACA05873 ID ACA05873 standard; cDNA; 3871 BP. XX ACA05873; XX	

QY 626 TACATTGAAATTGGACATCATCTGAGGCAATATATTTTAAGCTGGAGAACTCACATTTT 685
DB 575 TATATTGATCTTGGAGTCTCAAGAGAGATATATATTTTAACTAGACACATATTTTG 634
QY 686 GAAACCGCGTGGGAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGGCAGACCTGAAATGTCCTTTCGATCTCAGCAGCCTTTTTCAGTAAATGACA 694
QY 746 GATGGAGAAATTACTCTGGAATCTGAGCTGATTTTATGGGGGAGACCTTTGCTATCTTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTTCTTTGGCAAGATATCTGCATTTCACT 754
QY 806 CGAATCTTTGGGC-----ACCACCCCAATCAGGACAGAGCAGCATGATCCAGG 856
DB 755 CGATCCCTTTGGGCTACTCATGACCACTACTATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTGGCCACTCATCTCAGAGAGTGAACAATCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTTATGGAACCTTTCTTATACAGACACCTTACAATCCAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGGAATA 976
DB 875 GATGATAAATATATTTCTTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA 934
QY 977 GCTACTCACGCTAGAAATAGTGCAGATATGCAAGAAATGACTTTGGAGGGCACAGAACTCTG 1036
DB 935 ACCATCTTCTCGAGTGGAGAGTTTGTAGAAATGATGTAGAGGACACAGCGCCCTG 994
QY 1037 GTGAATPAATGGACAACTTCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAAT 1096
DB 995 ATAAACAAGTGGACGACTTTTCTTAAAGGCCAGACTGATTTGCTCAATCTCTGGAAGTAT 1054
QY 1097 GGCATTGACACTCATTTTCTGATGAACTGAGGATGATTTCTTAATGAACTTTAAAGATCT 1156
DB 1055 GGGCGAGATACCTTCTGATGAGCTTCAAGATATTTATTTACTCCCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGAGTGTATGACCTTCCAGTAAACATTTTCAAGGGATCAGCC 1216
DB 1115 AGNAATCTGTAGTATATGAGTCTTTTACTACACAGGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGATATAGCATGATGATGTAGAAGGGTGTCTTGGTCCATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGCTGACATCAGAGCAGTTTTTAAATGGTCCATATGCTCATAAG 1234
QY 1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCAGCGCCAGGA 1336
DB 1235 GAAAGTGCAGACCATCGTTGGGTGCGATGATGAGGAGAAATTCCTTATCCACGGCTGGT 1294
QY 1337 ACTTGTCCCGACAAAACATTTG---GTGGTGTGACTCTCAAAAGGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCGAAGCAAAACCTATGACCCACTGATTAAGTCCACCGGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAAT 1453
DB 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGGATGTAAATTTACAAATTTGCTGTAGACCGA 1513
DB 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTACAGACTGACACAGATAGTGTGGATCAT 1474
QY 1514 GTGGATGCAGAGATGGACAGTATGATGTATGTTTATCGGAACAGATGTTCGGGACCGTT 1573
DB 1475 GTCAATGCAAGAGATGGCCAGTACGATGTAAATGTTTCTTGGAAACAGACATTTGGAATGTC 1534
QY 1574 CTTAAAGTATGTTCAATCTTAAAGGAGACTTGGTATGATTTTAAAGAGAGTCTTCTGCTGAA 1633
DB 1535 CTTAAAGTGTGACGATTTTCAAAGGAAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTCTTGGGAAACCGACTGCTATTTTCAAGCAATGGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTGACATATTTCAAGCACTCATCAATCATCTTTGAAACATGGAATTTGTCTCTGAAGCAG 1651

QY 1694 CAACAACTATATATTTGGTTCAACGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGTGAT 1753
DB 1652 CACAAATTTGACATTTGGTTCCGAGATGATTTAGTTTCACTCTCTTGCACAGATGGAC 1711
QY 1754 ATTTACGGGAAGCGGTGTGCTGAGTGTGCTCGCCCGAGAGCCCTTACTGTGCTTGGGAT 1813
DB 1712 ACTTATGGGAAGCTTGGCAGACTGTGTGCTTGGCAGAGACCCCTACTGTGCTGGGAT 1771
QY 1814 GGTTCGCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCACAAAGACGACAAGATATA 1873
DB 1772 GGAATGATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGAGCGCAAGATGA 1831
QY 1874 AGAATGAGAGCCCACTCACTCACTGTTTCAAGCTTACACCATGATAATCACCATGGCCAC 1933
DB 1832 AAAATGCGGACCAATCACCAGTGTGGGACATCGAAGACAGCATTTAGTCTATG--AA 1888
QY 1934 AGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGAAATGCAAT 1993
DB 1889 ACTGCTGATGAAAAGGTGATTTTGGCAATTTAACTCAACCTTTCTGGAATGTATA 1948
QY 1994 CCGAAGTCGACAGAGCGCTGCTCTATTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 2053
DB 1949 CCTAATCCCAACAGCAACTATTAAATGGTATATCCAGAGTCAAGGATGAGCATCGA 2008
QY 2054 GAAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGT 2113
DB 2009 GAGGATTTGAAGCCCGATGAAAGAAATCATCAAAAGGAAATATGGGCTACTGATTCGAAT 2068
QY 2114 CTACAAAGAAAGGATTCAGGCAATTTACCTCTGCGAATGCGGTGGAAACATGGGTTTCAACNA 2173
DB 2069 TTGCAGAAAGAGGATTTCTGGGATGTATTCTGCAAAAGCCAGGAGCACACTTTTCTATCCAC 2128
QY 2174 ACTCTTCTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCATTTGGAAGAATCTTCTTCAT 2233
DB 2129 ACCATGTAAGCTGACTTTGAATGTCAATGAAGTGAACAGATGGAATAATACCCAGAGG 2188
QY 2234 AAAGATGATGATG 2246
DB 2189 GCAGAGCATGAGG 2201

RESULT 14

ABX97917

ID ABX97917 standard; cDNA; 3871 BP.

XX AC ABX97917;

XX DT 16-MAY-2003 (first entry)

XX Human PRO polynucleotide #174.

XX Human; PRO; gene; ss; cytostatic; chromosome mapping; gene mapping;
XX protein electrophoresis; tumour necrosis factor-alpha; TNF-alpha; blood;
XX chondrocyte differentiation; chondrocyte proliferation; tumour.

XX OS Homo sapiens.

XX FN US2003032102-A1.

XX PD 13-FEB-2003.

XX PF 17-JUN-2002; 2002US-0173697.

XX PR 16-SEP-1998; 98WO-US19330.

XX PR 07-OCT-1998; 98WO-US21141.

XX PR 01-DEC-1998; 98WO-US25108.

XX PR 08-MAR-1999; 99WO-US05028.

XX PR 14-MAY-1999; 99WO-US10733.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 01-SEP-1999; 99WO-US20111.

XX PR 15-SEP-1999; 99WO-US21090.

XX PR 01-DEC-1999; 99WO-US28301.

XX PR 02-DEC-1999; 99WO-US28551.

PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05941.
PR 15-MAR-2000; 2000WO-US06984.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083222P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.

PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.

PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 01-SEP-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098016P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.
Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206; Mismatches 739; Indels 21; Gaps 5;
Matches 1233; Conservative 0;

QY 269 AATGGAAGAACAAATGTGCCAAGGCTGAAATATCTCAAAAGAAATGTGGAATCCAAC 328
DB 215 ACTTGAAGCAAAATATCCAGACTCAAGCTAACCCTCAAGAGACTTGTCTCTCAAT 274
QY 329 AATGTGATCATTCAATGGCTTGGCCAAAGCTCAGTTCATACACTTCTCTTTGGAT 388
DB 275 AGCTGTATTCCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACTCTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTTGGAGCAAGAGATCACAATTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGGAGGCTGTCTTGGGAGCCAAAGACCAATCTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGTGCCAGTATCTTTACACCAAGAGAGATGAA 505
DB 395 GACTTAAACAAAAATTTTAAAGAAATTTATGGCTGTGCAAGGAGGCTGGAATTA 454
QY 506 TGCAGTGGCTGGAAAGACATCTCGAAAGAAATGTGCTTAATTTTCAATGAGTACTTAAG 565
DB 455 TGTAAATTAGCTGGGAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCAGCTTGTAGCCTGTGGAAGGGGGCTTTTCATCCAAATTTGCACC 625
DB 515 CCTATAACAAACTCACAATATATGTTGTGAGCTGAGAGCAATTCATCAATATGTGGG 574
QY 626 TACATTGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTT 685
DB 575 TATATTGATCTGGAGCTTACAAGGAGGATTTATATTCAAACTAGACACACATAATTG 634
QY 686 GAAACGGCTGGGAAGATGCAATATGACCTTAAGTCTGCTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTCGATCTCTCAGCAGCCTTTTGTCTTTCAGTAATGACA 694
QY 746 GATGAGAAATTAATCTCTGGAATCGAGCTGATTTTATGGGGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATACTGCATTCAT 754
QY 806 CGAACTCTTTGGG-----ACCACCAACCAATCAGGACAGAGCAGCATGATTTCCAGG 856
DB 755 CGATCCCTTTGGGCTACTCATGACCACTACATCAGAACTGACATTTTCAGAGCAGTAC 814
QY 857 TGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGACAACTCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTTATGGAATTTTCTTATACCAAGCAGCCTTACAAATCCAGAT 874

QY 917 GATGACAAAGTATACTTTTCTTCGTCGAAATGCAATAGATGGAGAAACACTCTCGAAAA 976
DB 875 GATGATAAAATATATTTCTTTCTGTAATCATCTCAAGAGGAGTACCTCCGATATA 934
QY 977 GCTACTCAGCTAGGAATAGGTACAGATATGCAAGAAATGACTTTGGAGGGCACAAGATCTG 1036
DB 935 ACCATCTCTCTCGAGTTGGAAGAGTTTGTAAAGATGATGTAGGAGGACAAAGCAGCTG 994
QY 1037 GTGAATAATGGACAAACATTTCTCAAGCTCGTCTGATTTCTCAGTCCAGTCCCAAT 1096
DB 995 ATAAACAAGTGGAGACTTTTCTTAAGGCCAGAGTGTGCTCAATTTCTTGAAGTGTAT 1054
QY 1097 GGCATTGACACTCATTTTGTGAACTGCAAGATGTATTCTTAATGAACTTTTAAAGATCT 1156
DB 1055 GGGCAGATATCTTACTTTGTAGAGCTTCAAGATATTTATTACTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGATATGGAGTGTTHAAGCTTCCAGTAAACATTTTCAAGGATCAGCC 1216
DB 1115 AGAAATCCTGTAGTATATGGAGTCTTTACTACAACAGCTCCATCTTCAAAGGCTCTGCT 1174
QY 1217 GTGTGTATAGCATGATGTGAGAGGGTGTCTTGTGTCATATGCCCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTAAATGGTCCATATGCTCATAG 1234
QY 1277 GATGACCCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCCATATCCAGCCGACGA 1336
DB 1235 GAAAGTCAGACCACTGTGTGGTGCAGTATGATGGAGAAATCTTATCCAGGCTGGT 1294
QY 1337 ACTTGTCCCAACAAAACATTTG---GTGGTTTGTACTTACAAGAGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCCAAACAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTCCAGCCAGTGTACATCCAGTGTTCCTTATGAAACAT 1453
DB 1355 GTCATCAGTTCTATAAAGCGCCTCTGTGTATATAAGTCCGTATATCCAGTTTCAGGA 1414
QY 1454 CGCCCAATAGTGTATCAAAACGAGTAAATATCAATTTACACAATTTGTGTAGACGA 1513
DB 1415 GGACCAAGTTCAAGAGATCAATGTGGNTTACAGCTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGATGACAGAGATGACAGTATGATTTATGTTTATCGAAACAGATGTTGGGACCTG 1573
DB 1475 GTCATGTCAGAAAGATGGCCAGTACGATGTAATTTTCTTGGAAACAGACATTTGGAAC 1534
QY 1574 CTTAAAGTGTCTCAATCTCAAGGAGTGTGATGATTTAGAGAGGTTCTGCTGGA 1633
DB 1535 CTAAAGTGTGAGCATTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGTCTATTTTCAAGCAATGGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTGCAGATATTCAGCACTCATCAATCATCTTGAACATGGAATTTCTCTGAGCAG 1651
QY 1694 CAACAACTATATATTTGTTCAACGGCTGGGTTGCCAGTCTCCCTTTTACACCGGTGTGAT 1753
DB 1652 CAACAAATGTACATTTGGTTCCCGAGATGGAATAGTTAGTTAGTCTCTCTTGCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCTCCCGAGACCTTACTGCTGTTGGAT 1813
DB 1712 ACTTATGGGAAGGTTGCGCAGACTGTTGTCTTCCAGAGACCCCTACTGTGCTGGGAT 1771
QY 1814 GGTTCGTGATCTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATA 1873
DB 1772 GGAATGCATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGA 1831
QY 1874 AGAAATGGAGACCACTGACTGCTGTTCAGCTTTACACCATGATATAATCACTGGCCAC 1933
DB 1832 AAATATGCGACCCAAATCACCAGTGTGCGACATCGAAGACAGCATTTAGTCAATG---AA 1888
QY 1934 AGCCCTGAGAGAGAAATCATCTATGTTGTAGAGAAATAGTACACATTTTTCGAAATGAGT 1993
DB 1889 ACTGCTGATGAAAGGTTGATTTTGGCAATGAAATTTTAACTCAACCTTTCTGGAATGTATA 1948
QY 1994 CCGAAGTCGAGAGAGCGCTGCTCTATTGGAATTTCCAGAGCGGCAATGAAAGCGCAAAA 2053

Db 1949 CTTAAATCCCAACAGCAACTATTAATGGTATATCCAGAGGTGAGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCCCTTCTGCTAGCTAGT 2113
Db 2009 GAGGAGTTGAGCCCGATGAAGATCATCAAAACGGAATATGGCTACTGATTCGAAGT 2068
Qy 2114 CTACACAGAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACATGGGTTTCATACA 2173
Db 2069 TTGCAGAGAAGGATTTCTGGGATGTTACTGCAAAAGCCAGGAGCACACTTTTCATCCAC 2128
Qy 2174 ACTCTCTTAAGSTACCTGGAAGTTCATTGACACAGAGCATTTGGAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGAATGTTCATTGAGAATGAACAGATGGAATAATACCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 15

ID ABX78701 standard; cDNA; 3871 BP.

AC ABX78701;

DT 15-APR-2003 (first entry)

DE Human PRO polynucleotide #174.

KW Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach;
KW liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.

OS Homo sapiens.

PN US2003027272-Al.

PD 06-FEB-2003.

PF 21-JUN-2002; 2002US-0176492.

XX 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.

PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 11-DEC-1997; 97US-066772P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078866P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 15-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.


```
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-097022P.
PR 18-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGACAATGTCGCAAGGCTGAAATTTATCTCAAAAGAAATGTTGGAATCCAAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTAACTCAAAAGACTTGTGCTTTCAAT 274
QY 329 AATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACATCTCTTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGACTGGATTTTCAAACTCTCTCTTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCAATTCAGTACCTGGTT 448
DB 335 GAGGAAGAGGCGAGCTGCTCTTGGGAGCCAAAGACCACATCTTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAGAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAAGAGATTTATTGGCCCTGCTGCAAGGAACGGGTGGAATTA 454
QY 506 TCGAAGTGGCTGGAAAAAGACATCTGAAAGAAATGTCTAAATTTTCATCAAGGTACTTAAG 565
DB 455 TGTAAATTTAGCTGGGAAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTT 514
QY 566 GCATATAATCAGACTCAGTGTACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACC 625
DB 515 CCTATAACAAACTCAGATATATGTGTGGAACTGGAGCATTTTCATCCAAATATGTGG 574
QY 626 TACATTGAAATTTGGACATCATCTTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTT 685
DB 575 TATATTGATCTTGGAGTCTCAAGAGGATATTTATTTCAAACTAGACACATATTTTG 634
QY 686 GAAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGCGACAGTGAATGTCTTTTCGATCCTCAGACGCTTTTGTCTTCAATATGACA 694
QY 746 GATGAGAAATTAATCTCTGGAACCTGACAGTGAATTTTATGGGGCGAGACTTGTCTATCTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTTCTTGGCAAGAGTACTGCAATCACT 754
QY 806 CGAATCTTTGGSC-----ACCACCAACCAATCAGACAGAGCAGCATGATCCAGG 856
DB 755 CGATCCCTTGGGCCCTACTCATGACCCACCATACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTGCACCCTCATCTCAGAGAGTGACAACTCGTAA 916
DB 815 TGGCTCAATGAGCAAAATTTATTGGAATTTTCTTCAACAGACACCTTACATCCAGAT 874
QY 917 GATGCAAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGGAGAGACACTCTGGAAA 976
DB 875 GATGATAAATATATTTCTTCTTCTGCGAATCATCTCAAGAAAGGCGAGTACCTCCGATAA 934
QY 977 GCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTGGAGGCGACAGAAGTCTG 1036
DB 935 ACCATCCTTCTCGAGTTGGAAGAGTTTGTAAAGATGATAGGAGGACACGCGACCTG 994
QY 1037 GTGAATAAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAAT 1096
DB 995 ATAAACAAGTGGACGACCTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGAAGTGT 1054
QY 1097 GGCATTGACATCAATTTTGATGAAGTGCAGAGATGATTTCTTAATGAACCTTTAAAGATCT 1156
DB 1055 GGGGAGATATCTTACTTTGATGAGCTTCAAGATATTTATTACTCTCCACAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGATATGAGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCC 1216
DB 1115 AGAAATCCTGATGATATGGAGTCTTTTACTTAACACAGCTCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGATAGCATGAGTGATGTGAGAAGGGTGTCTTGTGTCATATGCCACAGG 1276
```

Job time : 687 secs

```
Db 1175 GTTGTGTATAGCATGCTGACATCAGACAGTTTTAAATGGTCCATATGCTCATAG 1234
Qy 1277 GATGACCCCACTATCAATGGGTGCTTATCAAGAAAGAGTCCCTTATCCAGCGCAGGA 1336
Db 1235 GAAAGTGACACCATCGTTGGGTGAGTATGAGGAGATTCCTTATCCAGCGCTGGT 1294
Qy 1337 ACTTGTCCAGCAAAACATTG---GTGGTTTGTACTTCAAAAGACCTTCTCATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTTATGACCCACTGATTAAGTCCACCGAGATTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTCAGAGAGTCTATCCAGCCATGTACATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATCAGTTTCATAAGGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGA 1414
Qy 1454 CGCCCAATAGTGTCAAAACGGATGTAAATTTCAATTTACACAAATTTGCTAGACCGA 1513
Db 1415 GGACCAACGTTCAAGAGATCAATGTGGATTACAGACTGACACAGATAGTGGTGCAT 1474
Qy 1514 GTGGATGCAAGAGATGGACAGTATGATGTATGTATTCGGAACAGATGTGGGACCGTT 1573
Db 1475 GTCAATTGCAGAGATGGCCAGTACGATGTAATGTTTCTTGGAAACAGACATTCGAACTGTC 1534
Qy 1574 CTTAAGTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAA 1633
Db 1535 CTCAAAGTTGTGAGCATTTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTGGAG 1591
Qy 1634 GAAATGACAGTTTTTCGGGAAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCAG 1693
Db 1592 GAGTTGCAGATATTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTGAAGCAG 1651
Qy 1694 CAACAACTATATATTGGTTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTAT 1753
Db 1652 CAACAAATTGTACATTTGGTTCCCGAGATGGATTAGTTTCAGCTCTCCTTGCACAGATGCGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCCGAGACCCCTTACTGTGCTTGGGAT 1813
Db 1712 ACTTATGGGAAAGCTTGGCGACATGTTGTCTTGGCCAGAGACCCCTTACTGTGCTGGGAT 1771
Qy 1814 GGTCTGCTATGTTCTCGCTATTTTCCACTGCAAGAGACGCAAGACGACAGATATA 1873
Db 1772 GGAATGATGCTCTCGATATGCTCTTCTTCTTAAAGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGGAGACCCACTGACTCACTGTTTCAGACTTACACCATGATAATCACCATGGCCAC 1933
Db 1832 AAATATGGCGACCCAATCACCAGTGTGGGACATCGAAGACAGCATTAGTCATG---AA 1888
Qy 1934 AGCCCTGAAGACAGATCATCTATGTTGTAGAGATAGTAGACATTTTTTGGATGCGAGT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATA 1948
Qy 1994 CGAAGTGCAGAGAGCGTGTCTATTGGCAATTCAGAGGGGAAATGAAGAGCGMAAA 2053
Db 1949 CCTAATCCCAACAGCAACTATTAAATGGTATATCCAGAGGTAGGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGATGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGT 2113
Db 2009 GAGGAGTTGAAGCCCGATGAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT 2068
Qy 2114 CTACAACAGAGAGATTACGGCAATTTACCTCTGCCATGGGTGGAAATGGGTTTCATAAA 2173
Db 2069 TTGCAAGAAGAGGATTTCTGGGATGTATTACTGCAAAAGCCAGGAGCACATTTTCATCCAC 2128
Qy 2174 ACTCTTCTTAAGGTAAACCTGGAAGTCAATTGACACAGAGCATTTGGAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGAATGTCAATTGAGATGAACATGGAATAATACCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201
```

Search completed: July 31, 2003, 13:13:27

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:57:31 ; Search time 7224 Seconds
(without alignments)
12287.942 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatctttatttcgatg.....aggcttttttctctaataacc 2709

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:

1: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.oid.*
3: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US080 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US104A COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US104B COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6046 COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6047 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	33	US-09-774-490-1
2	2517	92.9	2848	48	US-10-247-671-83
3	2517	92.9	2848	71	US-60-172-373-8935
4	2517	92.9	2848	87	US-60-323-784-83

Sequence 1, Appli
Sequence 83, Appli
Sequence 8935, Ap
Sequence 83, Appli

```
5 2508.4 92.6 2530 1 PCT-US02-07826-283
6 2508.4 92.6 2530 15 US-09-053-375B-1350
7 2508.4 92.6 2530 45 US-10-097-340-283
8 2508.4 92.6 2530 49 US-10-262-538-9
9 2508.4 92.6 2601 9 US-08-488-935-1
10 2508.4 92.6 2601 9 US-08-489-057-1
11 2508.4 92.6 2601 45 US-10-067-632-53
12 2502.4 92.6 2601 47 US-10-170-235-19890
13 2050.4 92.6 2601 45 US-10-144-711-9178
14 2050.4 92.6 2601 45 US-10-144-711-9178
15 1544.8 75.7 2913 46 US-60-360-207-9178
16 1544.8 75.7 2913 91 US-09-649-164-8352
17 1544.8 75.7 2913 28 US-09-652-918-7308
18 1544.8 75.7 2913 28 US-09-652-918-7308
19 817 30.2 944 32 US-09-724-676-18078
20 718.6 30.2 944 32 US-09-724-676A-18078
21 718.6 26.5 2331 25 US-09-582-419A-2
22 718.6 26.5 2331 52 US-10-404-438-2
23 718.6 26.5 3871 40 US-09-946-374-309
24 718.6 26.5 3871 44 US-10-006-041A-309
25 718.6 26.5 3871 44 US-10-006-063A-309
26 718.6 26.5 3871 44 US-10-006-116A-309
27 718.6 26.5 3871 44 US-10-006-117A-309
28 718.6 26.5 3871 44 US-10-006-130A-309
29 718.6 26.5 3871 44 US-10-006-172A-309
30 718.6 26.5 3871 44 US-10-006-485A-309
31 718.6 26.5 3871 44 US-10-006-746A-309
32 718.6 26.5 3871 44 US-10-006-768A-309
33 718.6 26.5 3871 44 US-10-006-818A-309
34 718.6 26.5 3871 44 US-10-006-856A-309
35 718.6 26.5 3871 44 US-10-007-194A-309
36 718.6 26.5 3871 44 US-10-007-236A-309
37 718.6 26.5 3871 44 US-10-011-671A-309
38 718.6 26.5 3871 44 US-10-011-692A-309
39 718.6 26.5 3871 44 US-10-011-795B-309
40 718.6 26.5 3871 44 US-10-011-833A-309
41 718.6 26.5 3871 44 US-10-012-064A-309
42 718.6 26.5 3871 44 US-10-012-101B-309
43 718.6 26.5 3871 44 US-10-012-121A-309
44 718.6 26.5 3871 44 US-10-012-137A-309
45 718.6 26.5 3871 44 US-10-012-149A-309
```

ALIGNMENTS

```
RESULT 1
US-09-774-490-1
; Sequence 1, Application US/09774490
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1
```

```
Query Match 100.0%; Score 2709; DB 33; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTTATCGATGTTAAACAGCTTAGTAACTCGATCCAGGCTGAGGGTGTG 60
DB 1 AATCTTTTATCGATGTTAAACAGCTTAGTAACTCGATCCAGGCTGAGGGTGTG 60
```

```
QY 61 ACCACGCGTCCGGGAGTAGTGTAGCTCGCTGCTTCTCCATTGTCCAGCCAGTCTATTT 120
DB 61 ACCACGCGTCCGGGAGTAGTGTAGCTCGCTGCTTCTCCATTGTCCAGCCAGTCTATTT 120
QY 121 CCAGATTGTTGAATCTCTCTGGCGGCAATACAGGAAGGAGACTAAAGCAGCAAGG 180
DB 121 CCAGATTGTTGAATCTCTCTGGCGGCAATACAGGAAGGAGACTAAAGCAGCAAGG 180
QY 181 GACCTACAGCGTCTGACGATGGGCTGTAACTAGGATGCTCTCTTTCTTCTGGGGAGT 240
DB 181 GACCTACAGCGTCTGACGATGGGCTGTAACTAGGATGCTCTCTTTCTTCTGGGGAGT 240
QY 241 ATTACTACAGCAAGAGCAAACTATCAGAAATGGGAAGCAAAATGTCGAAGGCTGAAAT 300
DB 241 ATTACTACAGCAAGAGCAAACTATCAGAAATGGGAAGCAAAATGTCGAAGGCTGAAAT 300
QY 301 ATCTTACAAAAGAAATGTTGGAATCCAAATGTGTATCACTTTCAATGGCTGGCCCAACG 360
DB 301 ATCTTACAAAAGAAATGTTGGAATCCAAATGTGTATCACTTTCAATGGCTGGCCCAACG 360
QY 361 CTCCAGTTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAA 420
DB 361 CTCCAGTTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAA 420
QY 421 GGATCACATATTTCAATTCGACCTGTTTAAATATCAAGGATTTTCAAAAAGATTTGTGGCC 480
DB 421 GGATCACATATTTCAATTCGACCTGTTTAAATATCAAGGATTTTCAAAAAGATTTGTGGCC 480
QY 481 AGTATCTTACACAGAGAGATGAATCAATGTTGAGTCACTTTCAATGGCTGGCCCAACG 540
DB 481 AGTATCTTACACAGAGAGATGAATCAATGTTGAGTCACTTTCAATGGCTGGCCCAACG 540
QY 541 TGCTAAATTTCAATCAAGGTACTTAAAGGCATATAATCAGACTCACTTGTAGCCTGTGGAC 600
DB 541 TGCTAAATTTCAATCAAGGTACTTAAAGGCATATAATCAGACTCACTTGTAGCCTGTGGAC 600
QY 601 GGGGGCTTTTCAATCCAAATTTGACCTCAATTTGAAATGGAACATCATCTCGAGGACAAT 660
DB 601 GGGGGCTTTTCAATCCAAATTTGACCTCAATTTGAAATGGAACATCATCTCGAGGACAAT 660
QY 661 TTTTAAGCTGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAA 720
DB 661 TTTTAAGCTGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAA 720
QY 721 GCTGCTGACAGCTCCCTTTTAAATAGATGGAGAAATTAATCTGGAACCTGAGCTGATTT 780
DB 721 GCTGCTGACAGCTCCCTTTTAAATAGATGGAGAAATTAATCTGGAACCTGAGCTGATTT 780
QY 781 TATGGGCGGAGACTTTTGTCTATCTTCCGAACTCTTTGGGCACCACCACCAATCAGGACAG 840
DB 781 TATGGGCGGAGACTTTTGTCTATCTTCCGAACTCTTTGGGCACCACCACCAATCAGGACAG 840
QY 841 CGAGCATGATTTCCAGGTGGCTCAATGATCCAAAGTTCAATAGTGGCCCACTCATCTCAGA 900
DB 841 CGAGCATGATTTCCAGGTGGCTCAATGATCCAAAGTTCAATAGTGGCCCACTCATCTCAGA 900
QY 901 GAGTGACAACTCCTGAAGATGACAAAGTATATCTTTTCCGTAAGAAATGCAATAGATGG 960
DB 901 GAGTGACAACTCCTGAAGATGACAAAGTATATCTTTTCCGTAAGAAATGCAATAGATGG 960
QY 961 AGAACACTCTGGAAGAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGCAATGG 1020
DB 961 AGAACACTCTGGAAGAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGCAATGG 1020
QY 1021 AGGGCAGAGAGTCTGGTGAATAATGGAACAACTTCTCTCAAGCTCGTGTGATTTGCTC 1080
DB 1021 AGGGCAGAGAGTCTGGTGAATAATGGAACAACTTCTCTCAAGCTCGTGTGATTTGCTC 1080
QY 1081 AGTCCAGGTCCTCAATGCACTTCAATTTTGTAGTGAAGTGGAGGATGATTTCTTAAT 1140
DB 1081 AGTCCAGGTCCTCAATGCACTTCAATTTTGTAGTGAAGTGGAGGATGATTTCTTAAT 1140
QY 1141 GAACTTTAAAGATCCTAAAAATCCAGTTGTATATGGAGTGTTTAGGACTTCCAGTAACAT 1200
```

1141 GAACTTTAAAGATCTTAAATAATCCAGTGTGTATATGGAGTGTATAGCACTTCCAGTAACAT 1200
1201 TTTCAAGGATCAGCGGTGTATGTATAGCATGTAGTGTAGAGAGGTGTCTCTTGG 1260
1201 TTTCAAGGATCAGCGGTGTATGTATAGCATGTAGTGTAGAGAGGTGTCTCTTGG 1260
1261 TCCATATGCCCAACAGGGATGGACCAATCATCAATCAATGGGTGCTTATCAAGGAAGAGTCCC 1320
1261 TCCATATGCCCAACAGGGATGGACCAATCATCAATCAATGGGTGCTTATCAAGGAAGAGTCCC 1320
1321 CTATCCAGGCCAGGACCTTGTCCAGCAAAACATTTGGTGTGTGTGACTCTTACAAGGA 1380
1321 CTATCCAGGCCAGGACCTTGTCCAGCAAAACATTTGGTGTGTGTGACTCTTACAAGGA 1380
1381 CTTTCTGTATGTATTAACCTTTGCAAGAGTCAATCAGGCAATGTATCAATCCAGTGT 1440
1381 CTTTCTGTATGTATTAACCTTTGCAAGAGTCAATCAGGCAATGTATCAATCCAGTGT 1440
1441 TCTATGAACATCGCCCAATAGTATCAAAACCGGATGTAATTAATTAATTAACCAAT 1500
1441 TCTATGAACATCGCCCAATAGTATCAAAACCGGATGTAATTAATTAATTAACCAAT 1500
1501 TGTCTAGACCGAGTGGATGCAAGAGATGACAGTATGATGTATGTATGTTATCGGAACAGA 1560
1501 TGTCTAGACCGAGTGGATGCAAGAGATGACAGTATGATGTATGTATGTTATCGGAACAGA 1560
1561 TCTTGGGACCGTCTTAAAGTATGTTCAATTCCTAAGGAGACTTGTATGTTATGAAGA 1620
1561 TCTTGGGACCGTCTTAAAGTATGTTCAATTCCTAAGGAGACTTGTATGTTATGAAGA 1620
1621 GGTCTGCTGGAAGAAATGACAGTGTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCT 1680
1621 GGTCTGCTGGAAGAAATGACAGTGTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCT 1680
1681 TTCCTAAGCAGCAACACTATATATTTGGTTCAACGGTGGGTGGTGGCAGCTCCCTTT 1740
1681 TTCCTAAGCAGCAACACTATATATTTGGTTCAACGGTGGGTGGTGGCAGCTCCCTTT 1740
1741 ACACCGGTGTGATATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTTA 1800
1741 ACACCGGTGTGATATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTTA 1800
1801 CTGTCTTTGGATGTTCTGCAATGTTCTGCTATTTCCACTGCAAGAGACGCAAG 1860
1801 CTGTCTTTGGATGTTCTGCAATGTTCTGCTATTTCCACTGCAAGAGACGCAAG 1860
1861 AGCACAAGATATAAGAAATGGAGACCACTGACTCACTGTTTCCAGTGTATACCATGATAA 1920
1861 AGCACAAGATATAAGAAATGGAGACCACTGACTCACTGTTTCCAGTGTATACCATGATAA 1920
1921 TCACCATGGCCACAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTAGCACATT 1980
1921 TCACCATGGCCACAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTAGCACATT 1980
1981 TTTGGAATGCACTCCGAAGTCCAGAGAGCGTGTCTATTGGCAATTCAGAGCGGAAA 2040
1981 TTTGGAATGCACTCCGAAGTCCAGAGAGCGTGTCTATTGGCAATTCAGAGCGGAAA 2040
2041 TGAAGCGCAAAAGAGATCAGAGTGGATCATATCATCATCAGGACAGATCAAGGCT 2100
2041 TGAAGCGCAAAAGAGATCAGAGTGGATCATATCATCATCAGGACAGATCAAGGCT 2100
2101 TCTGTCTAGTGTCTTACAAGAGATTCAGGCAATACCTCTGCCATCGGTGGACA 2160
2101 TCTGTCTAGTGTCTTACAAGAGATTCAGGCAATACCTCTGCCATCGGTGGACA 2160
2161 TGGGTTTCATACAACTCTTTTAAAGTAACTTGAAGTCAATTCAGAGCATTTGGA 2220
2161 TGGGTTTCATACAACTCTTTTAAAGTAACTTGAAGTCAATTCAGAGCATTTGGA 2220
2221 AGAATCTTCTTAAAGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCGAATAG 2280

2221 AGAATCTTCTTAAAGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCGAATAG 2280
2281 CATGACACTAGCCAGAAAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACACCCCAA 2340
2281 CATGACACTAGCCAGAAAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACACCCCAA 2340
2341 TCTCAACAGATGATGATGTTCTGTGAACAAGTTTGGAAAGGGACCGAAAAACAAGTGG 2400
2341 TCTCAACAGATGATGATGTTCTGTGAACAAGTTTGGAAAGGGACCGAAAAACAAGTGG 2400
2401 GCAAGGCCAGGACATACCCAGGGAACAGTAACAATAATGGAAGCACTTACAAGAAAAATA 2460
2401 GCAAGGCCAGGACATACCCAGGGAACAGTAACAATAATGGAAGCACTTACAAGAAAAATA 2460
2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTTGAGAGGGCACCAGGAGTGTCTGAGCTGC 2520
2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTTGAGAGGGCACCAGGAGTGTCTGAGCTGC 2520
2521 ATTACCTTAGAAACCTCAACAGTAGAATACTTGCCTAGACAATACTTGGAACAAACA 2580
2521 ATTACCTTAGAAACCTCAACAGTAGAATACTTGCCTAGACAATACTTGGAACAAACA 2580
2581 TGCATATACATGAACTTTTTCATGCGCATTTATGTCGATTTTCAATGGTGGGAATTC 2640
2581 TGCATATACATGAACTTTTTCATGCGCATTTATGTCGATTTTCAATGGTGGGAATTC 2640
2641 AGCTGAGTTCACCAATTTATAATTTAAATTCATGAGTAACCTTCTTAATAGGCTTTTTT 2700
2641 AGCTGAGTTCACCAATTTATAATTTAAATTCATGAGTAACCTTCTTAATAGGCTTTTTT 2700
2701 CCTAATACC 2709
2701 CCTAATACC 2709

RESULT 2

US-10-247-671-83
; Sequence 83, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 83
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1930967CB1
US-10-247-671-83

Query Match 92.9%; Score 2517; DB 48; Length 2848;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGAGCATGGCTGGTTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTTACAGC 252
9 CTGAGCATGGCTGGTTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTTACAGC 68
QY 253 AAGGCAACTATCAGATGGGAACAATGTGCGAAGGCTGAATTTCTTACAAAGA 312
69 AAGGCAACTATCAGATGGGAACAATGTGCGAAGGCTGAATTTCTTACAAAGA 128
QY 313 AATGTTGAATCCCAACATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA 372

[illegible]

1209	Db	TGTTATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTTATGAACAA	1268
1453	QY	TCGCCCAATAGTGTACAAACCGGATGTAAATTTATCAATTTACACAAATTTGCTAGACCG	1512
1269	Db	TCGCCCAATAGTGTACAAACCGGATGTAAATTTATCAATTTACACAAATTTGCTAGACCG	1328
1513	QY	AGTGGATGCAGAAAGATGGACAGTATGATGTTATGTTTTATCGGAACAGATGTTGGGACCGT	1572
1329	Db	AGTGGATGCAGAAAGATGGACAGTATGATGTTATGTTTTATCGGAACAGATGTTGGGACCGT	1388
1573	QY	TCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA	1632
1389	Db	TCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA	1448
1633	QY	AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA	1692
1449	Db	AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA	1508
1693	QY	GCAACAACCTATATATTGGTTCAACCGCTCGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA	1752
1509	Db	GCAACAACCTATATATTGGTTCAACCGCTCGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA	1568
1753	QY	TATTTACGGGAAGCGTGTGAGTGTTCCTTCGCCCCGAGACCCCTTACTGTGCTTTGGGA	1812
1569	Db	TATTTACGGGAAGCGTGTGAGTGTTCCTTCGCCCCGAGACCCCTTACTGTGCTTTGGGA	1628
1813	QY	TGTTCTGCTGATGTTCTCGCTATTTTCCCACTCCAAAGAGACGCACAAAGACGACAAGATAT	1872
1629	Db	TGTTCTGCTGATGTTCTCGCTATTTTCCCACTCCAAAGAGACGCACAAAGACGACAAGATAT	1688
1873	QY	AAGAAATGGAGACCCACTGACTCACTGTTTCAGACTTTACACCATGATAATCACCATGGCCA	1932
1689	Db	AAGAAATGGAGACCCACTGACTCACTGTTTCAGACTTTACACCATGATAATCACCATGGCCA	1748
1933	QY	CAGCCCTGAAGAGAGAAATCATCTATGTTGTGTAGAGAAATGTAAGCAATTTTGGAAATGCAG	1992
1749	Db	CAGCCCTGAAGAGAGAAATCATCTATGTTGTGTAGAGAAATGTAAGCAATTTTGGAAATGCAG	1808
1993	QY	TCCGAAAGTCGACAGAGCGCTGCTTATTGGCAATTCACAGAGCGGAAATGAAGAGCGAAA	2052
1809	Db	TCCGAAAGTCGACAGAGCGCTGCTTATTGGCAATTCACAGAGCGGAAATGAAGAGCGAAA	1868
2053	QY	AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTACGTAG	2112
1869	Db	AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTACGTAG	1928
2113	QY	TCTACAACAGAGAGGATTCAGGCAATTAACCTCTGCCATCGGTTGGAACATGGTTTCATACA	2172
1929	Db	TCTACAACAGAGAGGATTCAGGCAATTAACCTCTGCCATCGGTTGGAACATGGTTTCATACA	1988
2173	QY	AACCTCTTTAAGGTAACCTCGAAGTCAATGACACAGAGCATTTTGGGAAGAACTTCTTCA	2232
1989	Db	AACCTCTTTAAGGTAACCTCGAAGTCAATGACACAGAGCATTTTGGGAAGAACTTCTTCA	2048
2233	QY	TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG	2292
2049	Db	TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG	2108
2293	QY	CCAGAAGGTCGTGTACAGAGACTTCATGACGTCAATCAACACCCCAATCTCTCAACAGAT	2352
2109	Db	CCAGAAGGTCGTGTACAGAGACTTCATGACGTCAATCAACACCCCAATCTCTCAACAGAT	2168
2353	QY	GGATGAGTTCTGTGACAAAGTTTGGBAAGGGACCGAACAACGTCGCGCAAGGCCAGG	2412
2169	Db	GGATGAGTTCTGTGACAAAGTTTGGBAAGGGACCGAACAACGTCGCGCAAGGCCAGG	2228
2413	QY	ACATATCCCGAGGAAACAGTAAACAAATTTGGAAGCACATTACAAGAAAAATGAAGAGGTAGAAA	2472
2229	Db	ACATATCCCGAGGAAACAGTAAACAAATTTGGAAGCACATTACAAGAAAAATGAAGAGGTAGAAA	2288
2473	QY	CAGGAGGACCCAGAAATTTGAGAGGCAACCCAGGAGTGTCTGAGCTGTGCATTCCTCTAGA	2532
2289	Db	CAGGAGGACCCAGAAATTTGAGAGGCAACCCAGGAGTGTCTGAGCTGTGCATTCCTCTAGA	2348


```
1569 TATTTACGGAAAGCGTGTCTGAGTGTTCCTCGCCGAGACCCCTACTGTCTGTGGGA 1628
1813 TGGTCTCGCATGTTCTCGTATATTTCCCACTCCAAAGACGACCAAGACGACGACATAT 1872
1629 TGGTCTCGCATGTTCTCGTATATTTCCCACTCCAAAGACGACCAAGACGACGACATAT 1688
1873 AAGAAATGAGAGACCCACTGACTCAGTCTGAGTGTTCAGACTTACACCATGATAATCACCATGGCCA 1932
1689 AAGAAATGAGAGACCCACTGACTCAGTCTGAGTGTTCAGACTTACACCATGATAATCACCATGGCCA 1748
1933 CAGCCCTGAGAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACACATTTTGGAAATGCGAG 1992
1749 CAGCCCTGAGAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACACATTTTGGAAATGCGAG 1808
1993 TCGGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGCGGAAATGAAGAGCGAAA 2052
1809 TCGGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGCGGAAATGAAGAGCGAAA 1868
2053 AAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 2112
1869 AAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 1928
2113 TCTAACACAGAGAGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAATCATGGTTTATACA 2172
1929 TCTAACACAGAGAGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAATCATGGTTTATACA 1988
2173 AACTCTTCTTAAGTTAAAGTCAAGTCAATGACACAGAGCAATTTGGAAAGAACTTCTTCA 2232
1989 AACTCTTCTTAAGTTAAAGTCAAGTCAATGACACAGAGCAATTTGGAAAGAACTTCTTCA 2048
2233 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2292
2049 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2108
2293 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2352
2109 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2168
2353 GGATGAGTCTGTGAACAAAGTTGGAAAGGAGGACCGAAACCAACGTCGCAAGGCCAGG 2412
2169 GGATGAGTCTGTGAACAAAGTTGGAAAGGAGGACCGAAACCAACGTCGCAAGGCCAGG 2228
2413 ACATACCCAGGAGAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2472
2229 ACATACCCAGGAGAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2288
2473 CAGGAGGACCCAGAAATTTGAGAGGACCCAGAGGTGTCTGAGCTGCATTTACCTTAGA 2532
2289 CAGGAGGACCCAGAAATTTGAGAGGACCCAGAGGTGTCTGAGCTGCATTTACCTTAGA 2348
2533 AACCTCAACAGTAGAAATTCCTCCTAGACATTAAGTGAAGAAACCAATGCAATATACAT 2592
2349 AACCTCAACAGTAGAAATTCCTCCTAGACATTAAGTGAAGAAACCAATGCAATATACAT 2408
2593 GAACTTTTTCATGGCAATTTGAGATGTTTCAATGTTGGAAATTCAGCTGAGTTTCA 2652
2409 GAACTTTTTCATGGCAATTTGAGATGTTTCAATGTTGGAAATTCAGCTGAGTTTCA 2468
2653 CCAATTAATAATTAATCATCAGTAACTTTCTTAATAGGCTTTTTTCTTAATACC 2709
2469 CCAATTAATAATTAATCATCAGTAACTTTCTTAATAGGCTTTTTTCTTAATACC 2525
```

RESULT 4

US-60-323-784-83

```
; Sequence 83, Application US/60323784
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
```

```
; FILE REFERENCE: PA-0050 P
; CURRENT APPLICATION NUMBER: US/60/323,784
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 83
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: Incyte ID No: 1930967CB1
US-60-323-784-83
```

Query Match 92.9%; Score 2517; DB 87; Length 2848;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 2517; Conservative 0;

```
QY 193 CTGCAGCATGGCTGGTTAACTAGGATTTCTGTCTTTTCTGGGAGATTAATTACTTACAGC 252
DB 9 CTGCAGCATGGCTGGTTAACTAGGATTTCTGTCTTTTCTGGGAGATTAATTACTTACAGC 68
QY 253 AAGACAAACTATCAGAAATGGAGAAACAATGTGCAAGGCTGAAATTAATCTTACAAAGA 312
DB 69 AAGACAAACTATCAGAAATGGAGAAACAATGTGCAAGGCTGAAATTAATCTTACAAAGA 128
QY 313 AATGTTGGAATCCAAACATGTGATCATTTCATTTGCTGGTGGCCACAGCTCCAGTTATCA 372
DB 129 AATGTTGGAATCCAAACATGTGATCATTTCATTTGCTGGTGGCCACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 432
DB 189 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 248
QY 433 TTCATTGCACTTAAATATCAAGGATTTTCAAAAGATTTGTGTCGACAGTATCTTACAC 492
DB 249 TTCATTGCACTTAAATATCAAGGATTTTCAAAAGATTTGTGTCGACAGTATCTTACAC 308
QY 493 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCA 552
DB 309 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCA 368
QY 553 CAAGTACTTAAAGGCATATAATCAGACTCAGTCTGTACGCTGTGGAACGGGGCTTTTCA 612
DB 369 CAAGTACTTAAAGGCATATAATCAGACTCAGTCTGTACGCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGACCTACATTTGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGACCTACATTTGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 488
QY 673 GAACTCAATTTTGAACACGGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
DB 489 GAACTCAATTTTGAACACGGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGGAATCTGAGCTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGGAATCTGAGCTGATTTTATGGGCGAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGAGCAGCATGATTC 852
DB 609 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGAGCAGCATGATTC 668
QY 853 CAGTGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGACATCC 912
DB 669 CAGTGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGACATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGAACACTCTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGAACACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAG 1032
DB 789 AAAAGCTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAG 848
```


Db 2109 CCAGAGGCTCTGTACAGAGACTTTCATGACGCTCATCAACACCCCAATCTCAACACGAT 2168
QY 2353 GGATGAGTTCTGTGAACAAAGTTTGAAAAGGACCGAAGAACCAACGTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTTCTGTGAACAAAGTTTGAAAAGGACCGAAGAACCAACGTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCCGAGGACAGTAAACAAATGAGACACTTACAGAAATAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGAGGACAGTAAACAAATGAGACACTTACAGAAATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGCAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGCAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2348
QY 2533 AACCTCAACAAAGTAGAAAATTGCTCTAGACAATAAATCTGGAAGAACCAATGCAATATACAT 2592
Db 2349 AACCTCAACAAAGTAGAAAATTGCTCTAGACAATAAATCTGGAAGAACCAATGCAATATACAT 2408
QY 2593 GAACTTTTTCATGCAATATGATGATGTTTACAAATGTTGGAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGCAATATGATGATGTTTACAAATGTTGGAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATATAAATTAATCATGAGTAACTTCTTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATATAAATTAATCATGAGTAACTTCTTAATAGGCTTTTTCCTCC 2518

RESULT 6

US-09-053-375B-1350

; Sequence 1350, Application US/09053375B

; GENERAL INFORMATION:

; APPLICANT: Chenchik, Alex

; APPLICANT: Bibilashvili, Robert

; TITLE OF INVENTION: Nucleic Acid Arrays

; FILE REFERENCE: CLON-006

; CURRENT APPLICATION NUMBER: US/09/053,375B

; CURRENT FILING DATE: 1998-08-31

; NUMBER OF SEQ ID NOS: 1543

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1350

; LENGTH: 2530

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-053-375B-1350

Query Match 92.6%; Score 2508.4; DB 15; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGACGATGGCTGGTTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTTACAGC 252
Db 9 CTGACGATGGCTGGTTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTTACAGC 68
QY 253 AAGACCAACTATCAGATGGGAAGAACAAATGTGCAAGGCTGAATATCTTACAAAGA 312
Db 69 AAGACCAACTATCAGATGGGAAGAACAAATGTGCAAGGCTGAATATCTTACAAAGA 128
QY 313 AATGTTGGAATCCAAATATGATCATCTTCAATGGCTTTGGCAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAAATATGATCATCTTCAATGGCTTTGGCAACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACAATTT 432
Db 189 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACAATTT 248
QY 433 TTCATTCGACCTGGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 492
Db 249 TTCATTCGACCTGGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 308
QY 493 CAGAAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGCTAAATTTTCAT 552
Db 309 CAGAAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGCTAAATTTTCAT 368

QY 553 CAAGGTACTTAAGGATATATATCAGACTCAGCTTGTACGCTTGTGAAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAGGATATATATCAGACTCAGCTTGTACGCTTGTGAAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTCGAGGACAAATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTCGAGGACAAATATTTTAAAGCTGA 488
QY 673 GAACTCAATTTTGAAGAACGGCCCGTGGGAAGAGTCCATATATACCCCTAAGCTGCTGACAGC 732
Db 489 GAACTCAATTTTGAAGAACGGCCCGTGGGAAGAGTCCATATATACCCCTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTAATAGATGAGAGATTTATCTCTGGAACCTGACAGCTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTTAATAGATGAGAGATTTATCTCTGGAACCTGACAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGCTATCTTCCGAACTCTTTGGGACACCCACCAATCAGGACAGAGCAGATGATTC 852
Db 609 CTTTGCTATCTTCCGAACTCTTTGGGACACCCACCAATCAGGACAGAGCAGATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATAGATGAGAGAACACTCTGG 788
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATAGATGAGAGAACACTCTGG 788
QY 973 AAAAGCTACTCACCTAGATAGGTGAGATATGCAAGATGACCTTTGGAGGGCAGAGAAG 1032
Db 789 AAAAGCTACTCACCTAGATAGGTGAGATATGCAAGATGACCTTTGGAGGGCAGAGAAG 848
QY 1033 TCTGTGAAATTAATGAGCAACATTTCTCAAGCTCGTCTGATTTGCTCAGTGGCAGGTCC 1092
Db 849 TCTGTGAAATTAATGAGCAACATTTCTCAAGCTCGTCTGATTTGCTCAGTGGCAGGTCC 908
QY 1093 AAATGGCAATGACACTCATTTTGTATGAACTGCGAGATGATATCTTAATGAATCTTTAAAGA 1152
Db 909 AAATGGCAATGACACTCATTTTGTATGAACTGCGAGATGATATCTTAATGAATCTTTAAAGA 968
QY 1153 TCCTAAAAATCCAGTGTATATGAGTGTGTTTACCACTTCCAGTAACTTTTCAAGGGATC 1212
Db 969 TCCTAAAAATCCAGTGTATATGAGTGTGTTTACCACTTCCAGTAACTTTTCAAGGGATC 1028
QY 1213 AGCGTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATGCCCA 1272
Db 1029 AGCGTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATGCCCA 1088
QY 1273 CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCC 1332
Db 1089 CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCC 1148
QY 1333 AGGAACCTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGAGGACCTTCTGATGA 1392
Db 1149 AGGAACCTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGAGGACCTTCTGATGA 1208
QY 1393 TGTATTAACCTTTTCAAGAAAGTCAATCCAGCATGTCAATCCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTTCAAGAAAGTCAATCCAGCATGTCAATCCAGTGTTCCTATGAACAA 1268
QY 1453 TCGCCCAATAGTATCAAAACGGATGTAAATTAATCAATTTACAAATTTGCTAGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACGGATGTAAATTAATCAATTTACAAATTTGCTAGACCG 1328
QY 1513 AGTGGATGCAAGAGATGAGATGATGTTTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGGATGCAAGAGATGAGATGATGTTTATGTTTATCGGAACAGATGTTGGACCGT 1388
QY 1573 TCTTAAAGTGTGTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGGA 1632
Db 1389 TCTTAAAGTGTGTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGGA 1448
QY 1633 AGAAATGACAGTGTTCCTGGGAACCGACTGCTATTTTACGAATGGAGCTTTTCCACTAAGCA 1692

QY 553 CAAAGTACTTAAGGCATATAATCAGACTCACTGTGTACGCTGTGGAAACGGGGGCTTTTCA 612
DB 369 CAAAGTACTTAAGGCATATAATCAGACTCACTGTGTACGCTGTGGAAACGGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACCTACATTTGAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACCTACATTTGAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 488
QY 673 GAACTCACATTTTGAACCGCGGTGGGAAGTCCATATGACCTTAAGCTGTCTGACAGC 732
DB 489 GAACTCACATTTTGAACCGCGGTGGGAAGTCCATATGACCTTAAGCTGTCTGACAGC 548
QY 733 ATCCCTTTTATAGATGAGAAATATATCTCTGGAATCTGAGCTGATTTTATGGGGCGAGA 792
DB 549 ATCCCTTTTATAGATGAGAAATATATCTCTGGAATCTGAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTCCTCTATCTCCGAACTCTTTGGGACACCAACCACTTCCAGGACAGAGCATGATTC 852
DB 609 CTTCCTCTATCTCCGAACTCTTTGGGACACCAACCACTTCCAGGACAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTTCGGTGAATAATGCAATAGATGAGAAACACTCTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTTCTTCGGTGAATAATGCAATAGATGAGAAACACTCTGG 788
QY 973 AAAAGCTACTCACCGCTAGAAATAGGTGAGATGCAAGATGAGTCTTGGAGGGCACAGAAG 1032
DB 789 AAAAGCTACTCACCGCTAGAAATAGGTGAGATGCAAGATGAGTCTTGGAGGGCACAGAAG 848
QY 1033 TCTGTGTGAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC 1092
DB 849 TCTGTGTGAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC 908
QY 1093 AAATGGCAATGACACTCACTTTTGTATGAACTGCAAGATGATTTCTTAATGAATTTTAAAGA 1152
DB 909 AAATGGCAATGACACTCACTTTTGTATGAACTGCAAGATGATTTCTTAATGAATTTTAAAGA 968
QY 1153 TCCTAAATCCAGTTGTATAGGTGATTTACGACTTCCAGTACATTTTCAAGGATC 1212
DB 969 TCCTAAATCCAGTTGTATAGGTGATTTACGACTTCCAGTACATTTTCAAGGATC 1028
QY 1213 AGCCGTGTGTATAGCATGAGTGTGTGAGAGGGTGTCTTGTGCTCATATGCCCA 1272
DB 1029 AGCCGTGTGTATAGCATGAGTGTGTGAGAGGGTGTCTTGTGCTCATATGCCCA 1088
QY 1273 CAGGATGGACCCAACTATCAATGGTGCCTTATCAAGGAAAGTCCCTTATCCACGGCC 1332
DB 1089 CAGGATGGACCCAACTATCAATGGTGCCTTATCAAGGAAAGTCCCTTATCCACGGCC 1148
QY 1333 AGGAATTTGTCACCAAAACATTTGGTGGTGTGACTCTCAAGAGGACCTTCTGTATGA 1392
DB 1149 AGGAATTTGTCACCAAAACATTTGGTGGTGTGACTCTCAAGAGGACCTTCTGTATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCACTCCAGGCAATGACAAATCCAGTGTCTTCTATGAACAA 1452
DB 1209 TGTATTAACCTTTGCAAGAGTCACTCCAGGCAATGACAAATCCAGTGTCTTCTATGAACAA 1268
QY 1453 TCGCCCAATAGTGTATCAAAACGGATGTAATTTATCAATTTTACAAAATGTCGTAGACCG 1512
DB 1269 TCGCCCAATAGTGTATCAAAACGGATGTAATTTATCAATTTTACAAAATGTCGTAGACCG 1328
QY 1513 AGTGATGCAAGATGACAGATGATGTTATGTTTATCGGAACAGATGTTGGGACCGT 1572
DB 1329 AGTGATGCAAGATGACAGATGATGTTATGTTTATCGGAACAGATGTTGGGACCGT 1388
QY 1573 TCTTAAAGTGTCTTCAATTTCTTAAGGACCTTGGTATGATTTTGAAGAGGTTCTGCTGGA 1632
DB 1389 TCTTAAAGTGTCTTCAATTTCTTAAGGACCTTGGTATGATTTTGAAGAGGTTCTGCTGGA 1448

QY 1633 AGAAATGACAGTTTTCGGGAACCGA CTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA 1692
DB 1449 AGAAATGACAGTTTTCGGGAACCGA CTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA 1508
QY 1693 GCAACAACTATATTTGGTTCAAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA 1752
DB 1509 GCAACAACTATATTTGGTTCAAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA 1568
QY 1753 TATTTACGGGAAACGGTGTGCTGAGTGTGGCTCGCCCGAGACCTTACTGTGCTTGGGA 1812
DB 1569 TATTTACGGGAAACGGTGTGCTGAGTGTGGCTCGCCCGAGACCTTACTGTGCTTGGGA 1628
QY 1813 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAAAGAGAGCAAGCAAGCAAGATAT 1872
DB 1629 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAAAGAGAGCAAGCAAGCAAGATAT 1688
QY 1873 AAGAAATGGACACCCACTGACTCACTGTTTCCAGACTTACACCATGATAATCACCATGGCCA 1932
DB 1689 AAGAAATGGACACCCACTGACTCACTGTTTCCAGACTTACACCATGATAATCACCATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAGAAATCATCTATGTTGAGAGAAATAGTAGCAATTTTGGAAATGCAG 1992
DB 1749 CAGCCCTGAAGAGAGAAATCATCTATGTTGAGAGAAATAGTAGCAATTTTGGAAATGCAG 1808
QY 1993 TCCGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGGGCAATGAAGAGCGAAA 2052
DB 1809 TCCGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGGGCAATGAAGAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 2112
DB 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 1928
QY 2113 TCTACAAACAGAGAGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTTCAACA 2172
DB 1929 TCTACAAACAGAGAGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTTCAACA 1988
QY 2173 AACTCTTCTTAAAGTAAACCTTGGAAAGTCAATTTGACACAGAGCATTTGGAAAGAACTTCTTCA 2232
DB 1989 AACTCTTCTTAAAGTAAACCTTGGAAAGTCAATTTGACACAGAGCATTTGGAAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGATGGAGATGGCTCTAGACCAAGAAATGTCCTAATAGCATGACACCTAG 2292
DB 2049 TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCTAATAGCATGACACCTAG 2108
QY 2293 CCAGAGGTCTGGTACAGAGACTTTCATGCACTCATCAACCAACCCCAATCTCAACAGCAT 2352
DB 2109 CCAGAGGTCTGGTACAGAGACTTTCATGCACTCATCAACCAACCCCAATCTCAACAGCAT 2168
QY 2353 GGATGAGTCTGTGAAACAAAGTTTGGAAAGGACCGGAAACAAACGTCGGCAAGGCCAGG 2412
DB 2169 GGATGAGTCTGTGAAACAAAGTTTGGAAAGGACCGGAAACAAACGTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGAAACAGTAACTTGAAGCACTTACAGAAATAGAAAGGTAGAAA 2472
DB 2229 ACATACCCAGGAAACAGTAACTTGAAGCACTTACAGAAATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGACCCACCAATTTTGAAGGACCGCCAGAGGTGTCTGAGCTGCATTACCTCTAGA 2532
DB 2289 CAGGAGACCCACCAATTTTGAAGGACCGCCAGAGGTGTCTGAGCTGCATTACCTCTAGA 2348
QY 2533 AACCTCAAAACAGTAGAAACTTTCCTAGACAAATTAACCTGGAAACAAATGCAATATACAT 2592
DB 2349 AACCTCAAAACAGTAGAAACTTTCCTAGACAAATTAACCTGGAAACAAATGCAATATACAT 2408
QY 2593 GAATCTTTTTCATGGCAATTTATGGGATGTTTACAAATGGTGGGAAATTCAGCTGAGTTCCA 2652
DB 2409 GAATCTTTTTCATGGCAATTTATGGGATGTTTACAAATGGTGGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTTAAATTAATCCCATGAGTAACTTCTTAAATAGGCTTTTTCCTCC 2702
DB 2469 CCAATTTAAATTAATCCCATGAGTAACTTCTTAAATAGGCTTTTTCCTCC 2518

RESULT 8

US-10-262-538-9
; Sequence 9, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)..(2331)
US-10-262-538-9

Query Match 92.6%; Score 2508.4; DB 49; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGGCTGTTAACTAGGATGCTCTCTTTCTGGGAGATTAATTACAGC	252
DB	9	CTGCAGCATGGGCTGTTAACTAGGATGCTCTCTTTCTGGGAGATTAATTACAGC	68
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGCTGAAATTAATCTCAAGA	312
DB	69	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGCTGAAATTAATCTCAAGA	128
QY	313	AATGTGGAAATCCAAATGTGATCACTTTCAATGCTTGGCCACAGCTCCAGTTATCA	372
DB	129	AATGTGGAAATCCAAATGTGATCACTTTCAATGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCTCTTTGGATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	432
DB	189	TACCTTCTCTTTGGATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	248
QY	433	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATGTTGTGGCCAGTATCTTAC	492
DB	249	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATGTTGTGGCCAGTATCTTAC	308
QY	493	CAGAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAGATGCTGTAATTTTCA	552
DB	309	CAGAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAGATGCTGTAATTTTCA	368
QY	553	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGTACGCTGTGGAAACGGGGCTTTTCA	612
DB	369	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGTACGCTGTGGAAACGGGGCTTTTCA	428
QY	613	TCCAATTCGACCTGAAATTTGGAACATCATCTCTGAGGACATATATTTTAAAGCTGA	672
DB	429	TCCAATTCGACCTGAAATTTGGAACATCATCTCTGAGGACATATATTTTAAAGCTGA	488
QY	673	GAACTACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTACCTGCTGACAGC	732
DB	489	GAACTACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTACCTGCTGACAGC	548
QY	733	ATCCCTTTTAAATAGATGGAGATTTATCTCTGAACTGAGCTGATTTTATGGGCGAGA	792
DB	549	ATCCCTTTTAAATAGATGGAGATTTATCTCTGAACTGAGCTGATTTTATGGGCGAGA	608
QY	793	CTTTGCTATCTTCGAACTCTTTGGGCAACCAACCAATCAGACAGAGCAGCATGATTC	852
DB	609	CTTTGCTATCTTCGAACTCTTTGGGCAACCAACCAATCAGACAGAGCAGCATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATCC	912
DB	669	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGG	972

DB	729	TGAAGATGACAAAGTATATCTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTCG	788
QY	973	AAAAGCTACTCACGCTAGAAATAGGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAG	1032
DB	789	AAAAGCTACTCACGCTAGAAATAGGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAG	848
QY	1033	TCTGTGTAATAAATGAGCAACAATCTCTCAAAAGCTGCTCTGATTTGCTCAGTCCAGGTC	1092
DB	849	TCTGTGTAATAAATGAGCAACAATCTCTCAAAAGCTGCTCTGATTTGCTCAGTCCAGGTC	908
QY	1093	AAATGGCAATTCACACTCAATTTTGATGAATCGAAGATGATTTCTTAATGAATCTTAAGA	1152
DB	909	AAATGGCAATTCACACTCAATTTTGATGAATCGAAGATGATTTCTTAATGAATCTTAAGA	968
QY	1153	TCCTAAAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTACATATTTCAAGGATC	1212
DB	969	TCCTAAAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTACATATTTCAAGGATC	1028
QY	1213	AGCCGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGGTCCATATGCCCA	1272
DB	1029	AGCCGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGGTCCATATGCCCA	1088
QY	1273	CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1148
QY	1333	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGGAGCTTCTGTATGA	1392
DB	1149	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGGAGCTTCTGTATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTGTCTTCTATGAA	1452
DB	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTGTCTTCTATGAA	1268
QY	1453	TCGCCCAATAGTGTATCAAAACCGATGTAAATTAATCAATTTACAAAATTTGCTAGACG	1512
DB	1269	TCGCCCAATAGTGTATCAAAACCGATGTAAATTAATCAATTTACAAAATTTGCTAGACG	1328
QY	1513	AGTGGATGCAAGATGAGACATGATGTATGTTATCGGACAGATGTTGGACCGT	1572
DB	1329	AGTGGATGCAAGATGAGACATGATGTATGTTATCGGACAGATGTTGGACCGT	1388
QY	1573	TCCTAAAGTACTTTCAATTTCTAAGGACATGTTGTTATGATTTAGAAAGTCTCTCTGA	1632
DB	1389	TCCTAAAGTACTTTCAATTTCTAAGGACATGTTGTTATGATTTAGAAAGTCTCTCTGA	1448
QY	1633	AGAAATGACAGTCTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTA	1692
DB	1449	AGAAATGACAGTCTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTA	1508
QY	1693	GCAACACTATATATTTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGA	1752
DB	1509	GCAACACTATATATTTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGA	1568
QY	1753	TATTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTTGGGA	1812
DB	1569	TATTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTTGGGA	1628
QY	1813	TGGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAAGACGAAGATAT	1872
DB	1629	TGGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAAGACGAAGATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCTGTTTCCAGCTTACCATGATGATCACCATGGCCA	1932
DB	1689	AAGAAATGGAGACCCACTGACTCTGTTTCCAGCTTACCATGATGATCACCATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAGATCATCTATGGTGTAGAGATAGTAGACATATTTTGGAAATGAG	1992
DB	1749	CAGCCCTGAAGAGAGATCATCTATGGTGTAGAGATAGTAGACATATTTTGGAAATGAG	1808
QY	1993	TCCAGATGCGAGAGAGCGCTGTTATGTTGCAATTTCCAGAGGCGCAATGAGAGCGAAA	2052

Db 1809 TCGAAGTCCGAGAGCGCTGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 1928
QY 2113 TCTACAACAGAGGATTCAGGCAATTCCTCTGCCATGCGGTGGAAACATGGGTTCATACA 2172
Db 1929 TCTACAACAGAGGATTCAGGCAATTCCTCTGCCATGCGGTGGAAACATGGGTTCATACA 1988
QY 2173 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCAATTCGGAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCAATTCGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
QY 2293 CCAGAGGTCTGGTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAGGTCTGGTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACACGAT 2168
QY 2353 GGATGAGTCTGTGAACAAGTTTGGAAAAGGACCGAAACAAAGCTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAAGTTTGGAAAAGGACCGAAACAAAGCTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGAAACAGTAACAATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGAAACAGTAACAATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGCCAGGAATTTGAGAGGCCACCGAGGAGTGTCTGAGTGCATTAACCTCTAGA 2532
Db 2289 CAGGAGGCCAGGAATTTGAGAGGCCACCGAGGAGTGTCTGAGTGCATTAACCTCTAGA 2348
QY 2533 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATACTGGAATAAAGAAAGGTAGAAA 2592
Db 2349 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATACTGGAATAAAGAAAGGTAGAAA 2408
QY 2593 GAACCTTTTTCATGGCAATATGTGGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGGCAATATGTGGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTAATAATTAATCATGAGTAACTTCTCTAATAGGCTTTTTTTTCC 2702
Db 2469 CCAATTAATAATTAATCATGAGTAACTTCTCTAATAGGCTTTTTTTTCC 2518

RESULT 9

US-08-488-935-1

; Sequence 1, Application US/08488935

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey S.

; APPLICANT: Kolodkin, Alex L.

; APPLICANT: Matthes, David

; APPLICANT: Bentley, David R.

; APPLICANT: O'Connor, Timothy

; TITLE OF INVENTION: Semaphorins

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.

; STREET: 268 Bush Street, Suite 3200

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,935

; FILING DATE:

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/489,057
; FILING DATE: 09-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLSCULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-488-935-1

Query Match

Best Local Similarity 92.6%; Score 2508.4; DB 9; Length 2601;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGACGATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTACAGC 252
Db 9 CTGACGATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTACAGC 68
QY 253 AAGACAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATATCTTACAAAGA 312
Db 69 AAGACAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATATCTTACAAAGA 128
QY 313 AATGTTGAAATCCAAACAATGTGATCACTTTCAATGGCTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGAAATCCAAACAATGTGATCACTTTCAATGGCTGGCCAAACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTGGATGAGGAACGAGTAGCTGTATCTTGGAGCAAAAGGATCAATATT 432
Db 189 TACCTTCCTTTGGATGAGGAACGAGTAGCTGTATCTTGGAGCAAAAGGATCAATATT 248
QY 433 TTCACTTCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 492
Db 249 TTCACTTCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTCAT 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTCAT 368
QY 553 CAAGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGACCTACATTTGAAATTTGGAACATCTCTGAGGACAAATATTTTAAAGCTGGA 672
Db 429 TCCAAATTTGACCTACATTTGAAATTTGGAACATCTCTGAGGACAAATATTTTAAAGCTGGA 488
QY 673 GAATCAATTTTGAAGAACGGCGGTGGGAAGAGTCCATATGACCTAAAGCTGTGACAGC 732
Db 489 GAATCAATTTTGAAGAACGGCGGTGGGAAGAGTCCATATGACCTAAAGCTGTGACAGC 548
QY 733 ATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGACCTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGACCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTTGGGCAACCAATCCCAATCAGGACAGAGCATGATTC 852
Db 609 CTTTGTCTATCTTCGAACTCTTTGGGCAACCAATCCCAATCAGGACAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCAITTAGTGGCCACCTCATCTCTCAGAGAGTGACAATCC 912

Db 669 CAGGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 728
Qy 913 TGAAGATGACAAAGTATACCTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATACCTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGG 788
Qy 973 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGGACAGAG 1032
Db 789 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGGACAGAG 848
Qy 1033 TCTGGTGAATAAATGACCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 1092
Db 849 TCTGGTGAATAAATGACCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 908
Qy 1093 AAATGCGATGACACTCACTNTTTCATGAACTGCAAGATGATTTCCCTAATGAACTTTAAAGA 1152
Db 909 AAATGCGATGACACTCACTNTTTCATGAACTGCAAGATGATTTCCCTAATGAACTTTAAAGA 968
Qy 1153 TCCTAAAAATCCAGTTGTATATGAGCTGTTTACGACTTCCAGTAACATTTTCAAGGGATC 1212
Db 969 TCCTAAAAATCCAGTTGTATATGAGCTGTTTACGACTTCCAGTAACATTTTCAAGGGATC 1028
Qy 1213 AGCCGTGTGTATGTATGATGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCA 1272
Db 1029 AGCCGTGTGTATGTATGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCA 1088
Qy 1273 CAGGGATGACCAACTATCAATGCGTGTCTTATCAAGAGAGTCCCTATCCAGGCC 1332
Db 1089 CAGGGATGACCAACTATCAATGCGTGTCTTATCAAGAGAGTCCCTATCCAGGCC 1148
Qy 1333 AGGAACCTGTCCAGCAAAACATTTGCTGTGTTTGAATCTCAAGAGGACTTCCCTGATGA 1392
Db 1149 AGGAACCTGTCCAGCAAAACATTTGCTGTGTTTGAATCTCAAGAGGACTTCCCTGATGA 1208
Qy 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAA 1268
Qy 1453 TCGCCCAATAGTATCAAAACGATGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1512
Db 1269 TCGCCCAATAGTATCAAAACGATGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1328
Qy 1513 AGTGGATGACAAAGATGGAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db 1329 AGTGGATGACAAAGATGGAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
Qy 1573 TCTTAAAGTATGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1632
Db 1389 TCTTAAAGTATGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1448
Qy 1633 AGAATGACAGTTTTCGGGAACCGACTGTATTTTCCAGCAATGGAGCTTTCCTAAGCA 1692
Db 1449 AGAATGACAGTTTTCGGGAACCGACTGTATTTTCCAGCAATGGAGCTTTCCTAAGCA 1508
Qy 1693 GCAACCACTATATATTTGTTTCAACGCTGCGGTTGCCAGCTCCCTTTACCGGTGTA 1752
Db 1509 GCAACCACTATATATTTGTTTCAACGCTGCGGTTGCCAGCTCCCTTTACCGGTGTA 1568
Qy 1753 TATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTACTGTCTTGGGA 1812
Db 1569 TATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTACTGTCTTGGGA 1628
Qy 1813 TGGTTCGATGTTCTCGTATTTTCCCACTGCAAGAGACCGCACAGACGACAGATAT 1872
Db 1629 TGGTTCGATGTTCTCGTATTTTCCCACTGCAAGAGACCGCACAGACGACAGATAT 1688
Qy 1873 AGAATGACAGACCGCACTGACTCACTGTTTCAAGCTTACACCATGATAATCACATGGCCA 1932
Db 1689 AGAATGACAGACCGCACTGACTCACTGTTTCAAGCTTACACCATGATAATCACATGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCGAG 1992

Db 1749 CAGCCCTGAAGAGAGAAATCATCTATGTGTAGAGAATAGTACACATTTTTCGAATGCGAG 1808
Qy 1933 TCCGAAGTCGAGAGAGAGCGCTGGTCTATTTGGCAATTTCCAGAGCGCAATTAAGAGCGAAA 2052
Db 1809 TCCGAAGTCGAGAGAGAGCGCTGGTCTATTTGGCAATTTCCAGAGCGCAATTAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGAAGAGCCCTTCTGCTACGTTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGAAGAGCCCTTCTGCTACGTTAG 1928
Qy 2113 TCTACACAGAGAGATTCAGGCAATTCCTGCTCCATGCGGTGGAGACATGGTTCATACA 2172
Db 1929 TCTACACAGAGAGATTCAGGCAATTCCTGCTCCATGCGGTGGAGACATGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGGTAACCTCGGAGTCAATGACACAGAGCAATTTGGAGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTCGGAGTCAATGACACAGAGCAATTTGGAGAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACCCCAATCTCAACAGAT 2168
Qy 2353 GATGAGTCTGTGAACAAAGTTTGGAAAGGAGCCGAAACAAAGCTCGGCAAGGCCAGG 2412
Db 2169 GATGAGTCTGTGAACAAAGTTTGGAAAGGAGCCGAAACAAAGCTCGGCAAGGCCAGG 2228
Qy 2413 ACATACCCCGAGGAAACAGTAACAAATGGAAGCACTTACAAAGAAATAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGAGGAAACAGTAACAAATGGAAGCACTTACAAAGAAATAAGAAAGGTAGAAA 2288
Qy 2473 CAGAGGACCCACCAAGATTTGAGAGGCAACCCAGAGTGTCTGAGCTGCATTAACCTCTAGA 2532
Db 2289 CAGAGGACCCACCAAGATTTGAGAGGCAACCCAGAGTGTCTGAGCTGCATTAACCTCTAGA 2348
Qy 2533 AACCTCAACAAAGTAGAACTTTCCTACAGCAATTAACCTGAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAACAAAGTAGAACTTTCCTACAGCAATTAACCTGAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGCGCAATTTATGAGTGTTCATGAGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGCGCAATTTATGAGTGTTCATGAGTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACCTTCCCTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACCTTCCCTAATAGGCTTTTTCCTCC 2518

RESULT 10
US-08-489-057-1
; Sequence 1, Application US/08489057

; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/489,057

FILING DATE: 09-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-2/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 16..2331

US-08-489-057-1

Query Match 92.6%; Score 2508.4; DB 9; Length 2601;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGAGCATGGCTGGTAACTAGGATGCTCTTTCTGGGAGTATTA	252
DB	9	CTGAGCATGGCTGGTAACTAGGATGCTCTTTCTGGGAGTATTA	68
QY	253	AAGAGCAACTATCAGATGGGAAGAAACAATGTGCCAGGCTGAAATATCTACAAAGA	312
DB	69	AAGAGCAACTATCAGATGGGAAGAAACAATGTGCCAGGCTGAAATATCTACAAAGA	128
QY	313	AATGTTGGATCCAAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA	372
DB	129	AATGTTGGATCCAAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA	188
QY	373	TACCTTCCTTTGGATGAGGACCGAGTAGGCTGTATGTTGGAGCAAGGATCAATATT	432
DB	189	TACCTTCCTTTGGATGAGGACCGAGTAGGCTGTATGTTGGAGCAAGGATCAATATT	248
QY	433	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTCA	552
DB	309	CAGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTCA	368
QY	553	CAAGTCTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGCTTTTCA	612
DB	369	CAAGTCTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGCTTTTCA	428
QY	613	TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	672
DB	429	TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	488
QY	673	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGC	732
DB	489	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGC	548
QY	733	ATCCCTTTTAAATGAGATGAGATTTACTCTGGAATCTGAGCTGATTTTATGGGGCGAGA	792
DB	549	ATCCCTTTTAAATGAGATGAGATTTACTCTGGAATCTGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTGCTATCTTCGAACTCTTTGGGACCAACCAACCAATCAGGACAGACGATGATTC	852
DB	609	CTTTGCTATCTTCGAACTCTTTGGGACCAACCAACCAATCAGGACAGACGATGATTC	668

QY	853	CAGTGGCTCAATGATCCAAAGTTCAATTAGTGCCCACTCATCTCAGAGAGTGACAATCC	912
DB	669	CAGTGGCTCAATGATCCAAAGTTCAATTAGTGCCCACTCATCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACTCTGG	972
DB	729	TGAAGATGACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACTCTGG	788
QY	973	AAAAGCTACTCACGCTAGATAGTGCAGATATGCAAGAAATGACTTTGGAGGSCACAGAAG	1032
DB	789	AAAAGCTACTCACGCTAGATAGTGCAGATATGCAAGAAATGACTTTGGAGGSCACAGAAG	848
QY	1033	TCTGGTGAATAAATGGACAAACATTCCTCAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC	1092
DB	849	TCTGGTGAATAAATGGACAAACATTCCTCAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC	908
QY	1093	AAATGGCAATGACACTCAATTTTGTGAACTGCAAGCTGATTTTCTTAATGAATTTAAAGA	1152
DB	909	AAATGGCAATGACACTCAATTTTGTGAACTGCAAGCTGATTTTCTTAATGAATTTAAAGA	968
QY	1153	TCCTAAAAATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAAACATTTTCAAGGATC	1212
DB	969	TCCTAAAAATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAAACATTTTCAAGGATC	1028
QY	1213	AGCCGTGTATGTATAGCATGATGTGAGAAAGGTGTTCCTTGGTCCATATGCCCA	1272
DB	1029	AGCCGTGTATGTATAGCATGATGTGAGAAAGGTGTTCCTTGGTCCATATGCCCA	1088
QY	1273	CAGGATGGACCCCAACTATCAATGGTCCCTTATCAAGGAAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGGACCCCAACTATCAATGGTCCCTTATCAAGGAAGTCCCTATCCAGGCC	1148
QY	1333	AGGAATCTTGTCCCAAGCAAAACATTTGGTGGTTTGTGACTCTCAAGAGGACCTTCTGATGA	1392
DB	1149	AGGAATCTTGTCCCAAGCAAAACATTTGGTGGTTTGTGACTCTCAAGAGGACCTTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCAATGATCAATCCAGTGTTCCTATGAACAA	1452
DB	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCAATGATCAATCCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCAATAGTGTCAAAACGGATGTAAATTAATCAATTTACAAAATGTCGTAGACCG	1512
DB	1269	TCGCCAATAGTGTCAAAACGGATGTAAATTAATCAATTTACAAAATGTCGTAGACCG	1328
QY	1513	AGTGGATGCAGAAGATGGACAGTATGATGTTATGTTATTCGGAACAGAGTGTGGACCGT	1572
DB	1329	AGTGGATGCAGAAGATGGACAGTATGATGTTATGTTATTCGGAACAGAGTGTGGACCGT	1388
QY	1573	TCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGA	1632
DB	1389	TCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGA	1448
QY	1633	AGAAATGACATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCA	1692
DB	1449	AGAAATGACATTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTA	1752
DB	1509	GCAACAACTATATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTA	1568
QY	1753	TATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGGA	1812
DB	1569	TATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGGA	1628
QY	1813	TGGTCTGCACTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACCAAGATAT	1872
DB	1629	TGGTCTGCACTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACCAAGATAT	1688
QY	1873	AAGAAATGGAGCCCACTGACTCTGTTTCAAGCTTACACCATGATTAATACCATGGCA	1932
DB	1689	AAGAAATGGAGCCCACTGACTCTGTTTCAAGCTTACACCATGATTAATACCATGGCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACATTTTGGAAATGCAG	1992

Db 1749 CAGCCCTGAAGAGAAATCATCTATGCTAGAGAATAGTACACATTTTGGAAATGCAG 1808
Qy 1993 TCCGAAGTCCGAGAGAGCCCTGGTCTATTGGCAATTCAGAGCGGCAATGAAGAGCGAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCCCTGGTCTATTGGCAATTCAGAGCGGCAATGAAGAGCGAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGGTAG 1928
Qy 2113 TCTACAACAGAGATTCAGGCAATTCCTCTGCCATCGGTGGAAATGCGGTTCATACA 2172
Db 1929 TCTACAACAGAGATTCAGGCAATTCCTCTGCCATCGGTGGAAATGCGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGTGAACCTGGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTGAACCTGGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGAGTGGCTTAAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGAGTGGCTTAAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCAAGGCCAGG 2228
Qy 2413 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAA 2472
Db 2229 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAA 2288
Qy 2473 CAGGAGGCCAGCAATTTGAGAGGCCACCCAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2532
Db 2289 CAGGAGGCCAGCAATTTGAGAGGCCACCCAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2348
Qy 2533 AACCTCAACAGTAGAACAATTCGCTAGACATACTGAAACCAATGCAATATACAT 2592
Db 2349 AACCTCAACAGTAGAACAATTCGCTAGACATACTGAAACCAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCAATTTATGGAATTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCAATTTATGGAATTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTTCC 2702
Db 2469 CCAATTATAAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTTCC 2518

RESULT 11
US-10-067-632-53
; Sequence 53, Application US/10067632
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; Kolodkin, Alex L.
; Matthes, David
; Bentley, David R.
; O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/067,632
FILING DATE: 04-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/835,268
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-067-632-53

Query Match 92.6%; Score 2508.4; DB 45; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 193 CTGCAGCATGGCTGGTTAACTAGGATTGTCTGCTTTTCTGGGAGTATTACTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGGATTGTCTGCTTTTCTGGGAGTATTACTTACAGC 68
Qy 253 AAGAGCAACTATCAGAATGGAGAAACAATGTGCCAAGGCTGAATTTATCTACAAAGA 312
Db 69 AAGAGCAACTATCAGAATGGAGAAACAATGTGCCAAGGCTGAATTTATCTACAAAGA 128
Qy 313 AATGTTGGAATCCCAACATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCCAACATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
Qy 373 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 432
Db 189 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 248
Qy 433 TTCATTGCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 492
Db 249 TTCATTGCACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 308
Qy 493 CAGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAATTTTCA 552
Db 309 CAGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAATTTTCA 368
Qy 553 CAAGGTACTTAAGGCATATAATCAGACTCCTGTTACGCTGTGTGAAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAGGCATATAATCAGACTCCTGTTACGCTGTGTGAAACGGGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACATATTTTAAAGCTGA 488
Qy 673 GAACTCACATTTTGAACCGCCGCTGGGAAGAGTCCATATGACCCCTAGCTCTGACAGC 732
Db 489 GAACTCACATTTTGAACCGCCGCTGGGAAGAGTCCATATGACCCCTAGCTCTGACAGC 548
Qy 733 ATCCCTTTTAATAGATGGAGAAATTATATCTCTGGAATCTGAGCTGATTTTATGGGGCGAGA 792

i ORGANISM: HUMAN
US-10-170-235-19890

Query Match		92.4%;	Score 2502.4;	DB 47;	Length 2857;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 2516;		Conservative 0;	Mismatches 1;	Indels 3;	Gaps 1;
QY	193	CTGCAGCATGGGCTGGTAACTAGGATGTCTGTCTTTCTCGGGAGTAGTACTTACAGC	252		
DB	1	CTGCAGCATGGGCTGGTAACTAGGATGTCTGTCTTTCTCGGGAGTAGTACTTACAGC	60		
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTCAAAATTAATCTTACAAAGA	312		
DB	61	AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTCAAAATTAATCTTACAAAGA	120		
QY	313	AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCACACAGCTCCAGTTATCA	372		
DB	121	AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCACACAGCTCCAGTTATCA	180		
QY	373	TACCTTCCTTTTGGATGAGGACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT	432		
DB	181	TACCTTCCTTTTGGATGAGGACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT	240		
QY	433	TTCAATTCGACCTGGTTAATPATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	492		
DB	241	TTCAATTCGACCTGGTTAATPATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	300		
QY	493	CAGAGAGATGAATGCAAGTGGCTGGGAGGAGTGGTATGTTGGAGCAAAAGGATGTCTTAATTTCA	552		
DB	301	CAGAGAGATGAATGCAAGTGGCTGGGAGGAGTGGTATGTTGGAGCAAAAGGATGTCTTAATTTCA	360		
QY	553	CAAGGTACTTAAGGCATATAATCAGACTCACCTTTGTACGCTGTGGAACGGGGGCTTTTCA	612		
DB	361	CAAGGTACTTAAGGCATATAATCAGACTCACCTTTGTACGCTGTGGAACGGGGGCTTTTCA	420		
QY	613	TCCAAATTCGACCTAATGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGA	672		
DB	421	TCCAAATTCGACCTAATGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGA	480		
QY	673	GAACTCACATTTTGAACCGGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	732		
DB	481	GAACTCACATTTTGAACCGGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	540		
QY	733	ATCCCTTTTAAATAGATGGAGAAATTAATCTGGAACCTGAGCTGATTTTATGGGCGAGA	792		
DB	541	ATCCCTTTTAAATAGATGGAGAAATTAATCTGGAACCTGAGCTGATTTTATGGGCGAGA	600		
QY	793	CTTTGCTATCTTCCGAACTCTTTGGCACCAACCACCAATCAGGACAGACGATGATTC	852		
DB	601	CTTTGCTATCTTCCGAACTCTTTGGCACCAACCACCAATCAGGACAGACGATGATTC	660		
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCC	912		
DB	661	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCC	720		
QY	913	TGAAGATGACAAAGTATACCTTTCTCGTGAATGCAATAGATGAGAGACACTCTGG	972		
DB	721	TGAAGATGACAAAGTATACCTTTCTCGTGAATGCAATAGATGAGAGACACTCTGG	780		
QY	973	AAAAAGCTACTACGCTAGAAATAGGTGAGATATGCAAGATGATCTTTGGAGGCGACAGAAG	1032		
DB	781	AAAAAGCTACTACGCTAGAAATAGGTGAGATATGCAAGATGATCTTTGGAGGCGACAGAAG	840		
QY	1033	TCTGGTGAATAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTTCC	1092		
DB	841	TCTGGTGAATAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTTCC	900		
QY	1093	AAATGGCATTTGACCTCACTTTTGAATGAACTGAGGATGATTTCTTAAATGAACTTTAAAGA	1152		
DB	901	AAATGGCATTTGACCTCACTTTTGAATGAACTGAGGATGATTTCTTAAATGAACTTTAAAGA	960		
QY	1153	TCCTAAAAATCCAGTTGTATGGAAGTGTTCAGACTTCCAGTAACTTTTCAAGGGATC	1212		

DB	961	TCCTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAACTTTTCAAGGGATC	1020
QY	1213	AGCGGTGTATGTATAGCATGATGTAGAGAGGTGTTCCTTGGTCCATATGCCCA	1272
DB	1021	AGCGGTGTATGTATAGCATGATGTAGAGAGGTGTTCCTTGGTCCATATGCCCA	1080
QY	1273	CAGGATCGGACCCAACTACTCAATGGTGCCTTTATCAAGGAAGAGTCCCTATCCAGGCG	1332
DB	1081	CAGGATCGGACCCAACTACTCAATGGTGCCTTTATCAAGGAAGAGTCCCTATCCAGGCG	1140
QY	1333	AGGAATCTTGTCCAGCAAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGATGA	1392
DB	1141	AGGAATCTTGTCCAGCAAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGATGA	1200
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTTCCTATGAACAA	1452
DB	1201	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTTCCTATGAACAA	1260
QY	1453	TCGCCCAATAGTGTCAAAACGGATGTAAATTTACAAATTTACAAATTTGTCTAGACCG	1512
DB	1261	TCGCCCAATAGTGTCAAAACGGATGTAAATTTACAAATTTGTCTAGACCG	1320
QY	1513	AGTGGATCGAAGATGACAGTATGATGTTATGTTATCGGAAACAGATGTTGGACCGT	1572
DB	1321	AGTGGATCGAAGATGACAGTATGATGTTATGTTATCGGAAACAGATGTTGGACCGT	1380
QY	1573	TCCTAAAGTACTTCAATTTCCCTAAGGACITGGTATGATTTAGAGAGGTCTCTCTGA	1632
DB	1381	TCCTAAAGTACTTCAATTTCCCTAAGGACITGGTATGATTTAGAGAGGTCTCTCTGA	1440
QY	1633	AGAAATGACAGTTCCTCGGAAACCGACTGTCTATTTTACCAATTTGAGAGCTTTTCCACTA	1689
DB	1441	AGAAATGACAGTTCCTCGGAAACCGACTGTCTATTTTACCAATTTGAGAGCTTTTCCACTA	1500
QY	1690	CGAGCAACAACTATATATTGTTTCAACGGCTGGGGTTCGCCAGCTCTTACACGGTG	1749
DB	1501	CGAGCAACAACTATATATTGTTTCAACGGCTGGGGTTCGCCAGCTCTTACACGGTG	1560
QY	1750	TGATATTTACGGGAAAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGTGCTTG	1809
DB	1561	TGATATTTACGGGAAAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGTGCTTG	1620
QY	1810	GGATGGTCTCATGTTCTCGCTATTTTCCCACTGCAAAAGAGAGCGCAAGAACAAGA	1869
DB	1621	GGATGGTCTCATGTTCTCGCTATTTTCCCACTGCAAAAGAGAGCGCAAGAACAAGA	1680
QY	1870	TATAAGAAATGGAGACCCACTGACTCTGTTTACAGACTTACCAATGATTAATCACCATGG	1929
DB	1681	TATAAGAAATGGAGACCCACTGACTCTGTTTACAGACTTACCAATGATTAATCACCATGG	1740
QY	1930	CCAGAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCATTCTTGGATG	1989
DB	1741	CCAGAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCATTCTTGGATG	1800
QY	1990	CAGTCCGAAGTCGAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGCGGCAATGAAGAGCG	2049
DB	1801	CAGTCCGAAGTCGAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGCGGCAATGAAGAGCG	1860
QY	2050	AAAAAGAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACG	2109
DB	1861	AAAAAGAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACG	1920
QY	2110	TAGTCTACACAGAGAGGATTCAGGCAATTAACCTCTGCCATCGGTTGGAAACATGGGTTCT	2169
DB	1921	TAGTCTACACAGAGAGGATTCAGGCAATTAACCTCTGCCATCGGTTGGAAACATGGGTTCT	1980
QY	2170	ACAAACTCTTCTTAAGGTAAACCTCGGAAGTCAITTGACACAGAGCATTTTGGAGAACTTCT	2229
DB	1981	ACAAACTCTTCTTAAGGTAAACCTCGGAAGTCAITTGACACAGAGCATTTTGGAGAACTTCT	2040
QY	2230	TCATAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACC	2289
DB	2041	TCATAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACC	2100

Db	689	CTT	CAT	GGG	ACG	GACTT	CGCT	TAT	CTT	CAG	AAC	ACT	GGG	GAC	CAT	CAC	CCC	AT	CAG	GAC	748
Qy	838	AG	AC	G	A	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	897
Db	749	GG	AC	G	A	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	808
Qy	898	AG	AC	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	957
Db	809	AG	AC	G	A	T	G	A	T	G	A	T	G	A	T	G	A	T	G	A	868
Qy	958	TG	GA	GA	AC	ACT	CT	CG	AA	AG	CT	CT	CA	CG	T	AG	AA	T	AG	GT	1017
Db	869	CG	GA	GA	AC	ACT	CT	CG	AA	AG	CT	CT	CA	CG	T	AG	AA	T	AG	GT	928
Qy	1018	TG	GA	GG	CA	C	A	G	A	G	T	CT	GT	GA	TA	AA	T	AG	CA	AA	1077
Db	929	TG	GT	GG	CA	C	A	G	A	G	T	CT	GT	GA	TA	AA	T	AG	CA	AA	988
Qy	1078	CT	CA	G	T	G	C	A	G	G	T	CC	AA	T	GC	AT	G	A	CT	CA	1137
Db	989	CT	CT	GT	GC	CG	GT	CC	CA	T	G	CA	T	G	CA	T	G	CA	T	G	1048
Qy	1138	AA	T	GA	A	CT	T	T	AA	AG	AT	CC	T	AA	AA	T	CC	AG	AT	CC	1197
Db	1049	AA	T	GA	A	CT	T	T	AA	AG	AT	CC	T	AA	AA	T	CC	AG	AT	CC	1108
Qy	1198	CA	T	T	T	CA	A	G	G	G	A	T	C	A	G	T	G	A	T	G	1257
Db	1109	CA	T	C	T	T	AA	G	G	G	A	T	C	A	G	T	G	A	T	G	1168
Qy	1258	TG	GT	CC	AT	A	T	G	C	A	T	G	C	A	T	G	C	A	T	G	1317
Db	1169	TG	GT	CC	AT	A	T	G	C	A	T	G	C	A	T	G	C	A	T	G	1228
Qy	1318	CC	C	T	A	T	CC	AC	GG	CA	G	A	T	GT	CC	CA	G	A	A	A	1377
Db	1329	CC	C	T	A	T	CC	AC	GG	CA	G	A	T	GT	CC	CA	G	A	A	A	1288
Qy	1378	GG	AC	C	T	T	CC	T	G	A	T	G	T	T	AA	C	T	T	GA	A	1437
Db	1289	GG	AC	C	T	T	CC	T	G	A	T	G	T	T	AA	C	T	T	GA	A	1348
Qy	1438	GT	T	CC	T	A	T	GA	CA	AT	GC	CC	AA	T	GA	T	G	A	T	G	1497
Db	1349	GT	T	CC	T	A	T	GA	CA	AT	GC	CC	AA	T	GA	T	G	A	T	G	1408
Qy	1498	AA	T	T	G	T	CG	T	AG	AC	CG	AT	GC	CA	AG	AT	GC	CA	AG	AT	1557
Db	1409	AA	T	T	G	T	CG	T	AG	AC	CG	AT	GC	CA	AG	AT	GC	CA	AG	AT	1468
Qy	1558	AG	A	T	G	T	T	GG	AC	CG	T	T	CA	A	T	T	CC	T	AA	G	1617
Db	1469	AG	A	T	G	T	T	GG	AC	CG	T	T	CA	A	T	T	CC	T	AA	G	1528
Qy	1618	AG	AG	G	T	T	CG	T	GA	GA	AA	T	GA	C	A	G	T	T	T	T	1677
Db	1529	AG	AG	G	T	T	CG	T	GA	GA	AA	T	GA	C	A	G	T	T	T	T	1588
Qy	1678	GC	T	T	T	CC	A	T	AG	CA	CA	A	C	T	A	T	A	T	T	T	1737
Db	1589	GC	T	T	T	CC	A	T	AG	CA	CA	A	C	T	A	T	A	T	T	T	

```

RESULT 15
US-09-649-164-8352
; Sequence 8352, Application US/09649164
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1165.001
; CURRENT APPLICATION NUMBER: US/09/649,164
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/150,606
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 9995

```

FASTSEQ for Windows Version 4.0
; SEQ ID NO 8352
; LENGTH: 2256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)---(2256)
; OTHER INFORMATION: n = A,T,C or G
US-09-649-164-8352

Query Match 57.0%; Score 1544.8; DB 28; Length 2256;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1546; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	59	CGACCCACCGCTCGGAGTAGTTGAGTTCGCTCGCTGTTCTCCCATGTGACGACGCTCTAT	118
DB	1	CGACCCACCGCTCGGAGTAGTTGAGTTCGCTCGCTGTTCTCCCATGTGACGACGCTCTAT	60
QY	119	TTCCAGATTGTTGAACTTCTCTGGCCGACAACTATCAGGAAGGAAGACTTAAAGCAGCAAA	178
DB	61	TTCCAGATTGTTGAACTTCTCTGGCCGACAACTATCAGGAAGGAAGACTTAAAGCAGCAAA	120
QY	179	GGGACCTACAGCTCTGACGATGGCTGTTAACTAGGATTGTCTGTTTCTGGGA	238
DB	121	GGGACCTACAGCTCTGACGATGGCTGTTAACTAGGATTGTCTGTTTCTGGGA	180
QY	239	GTATTACTTACAGCAGCAAACTATCAGATGGGAAGCAAACTGTGCCAAGGCTGAAA	298
DB	181	GTATTACTTACAGCAGCAAACTATCAGATGGGAAGCAAACTGTGCCAAGGCTGAAA	240
QY	299	TTATCTCTACAAAGAAATGTTGAATCCAACTATGATCACTTTCAATGGCTTGGCCCAAC	358
DB	241	TTATCTCTACAAAGAAATGTTGAATCCAACTATGATCACTTTCAATGGCTTGGCCCAAC	300
QY	359	AGCTCCAGTTATCATACCTTCTTTGGATGAGGAACGGAGTAGCTGTATGTTGGACA	418
DB	301	AGCTCCAGTTATCATACCTTCTTTGGATGAGGAACGGAGTAGCTGTATGTTGGACA	360
QY	419	AGGATCATATTTTCATCTGACCTGTTTATATCAAGATTTTCAAAAGATTGTGTGG	478
DB	361	AGGATCATATTTTCATCTGACCTGTTTATATCAAGATTTTCAAAAGATTGTGTGG	420
QY	479	CAAGTATCTTACACCAAGAGAGATGAATGCAAGTGGCTGGAAGAGACATCTGAAAGAA	538
DB	421	CAAGTATCTTACACCAAGAGAGATGAATGCAAGTGGCTGGAAGAGACATCTGAAAGAA	480
QY	539	TGTGCTAATTTTCATCAAGTACTTAAAGCATATATCAGACTCACTTTGACGCTGTGGA	598
DB	481	TGTGCTAATTTTCATCAAGTACTTAAAGCATATATCAGACTCACTTTGACGCTGTGGA	540
QY	599	ACGGGGCTTTTCATCCAAATTTGACCTTACATTTGAATTTGGACATCATCTGAGGACAAT	658
DB	541	ACGGGGCTTTTCATCCAAATTTGACCTTACATTTGAATTTGGACATCATCTGAGGACAAT	600
QY	659	ATTTTAAAGCTGAGAACTCACATTTTGAAGCGGCGTGGGAAGAGTCCATATGACCT	718
DB	601	ATTTTAAAGCTGAGAACTCACATTTTGAAGCGGCGTGGGAAGAGTCCATATGACCT	660
QY	719	AAGCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAATGACGCTGAT	778
DB	661	AAGCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAATGACGCTGAT	720
QY	779	TTTATGGGCGAGACTTTGCTATCTTCGAATCTTTGGGCACCAACCCCAATCAGGACA	838
DB	721	TTTATGGGCGAGACTTTGCTATCTTCGAATCTTTGGGCACCAACCCCAATCAGGACA	780
QY	839	GAGCAGCATGATTCAGGTGGCTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATCTCA	898
DB	781	GAGCAGCATGATTCAGGTGGCTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATCTCA	840
QY	899	GAGAGTGAACAATCTCTGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGAT	958

DB	841	GAGAGTGAACAATCTCTGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGAT	900
QY	959	GGAGAACACTCTGGAAAAGCTACTCAGCTAGAATAGTCAAGATATGCAAGAACTGACCTTT	1018
DB	901	GGAGAACACTCTGGAAAAGCTACTCAGCTAGAATAGTCAAGATATGCAAGAACTGACCTTT	960
QY	1019	GGAGGGCACAGAGTCTGGTGAATAAATGGAACAACATTTCTCAAAGCTCGTCTGATTTGC	1078
DB	961	GGAGGGCACAGAGTCTGGTGAATAAATGGAACAACATTTCTCAAAGCTCGTCTGATTTGC	1020
QY	1079	TCAGTGCAGGCTCAAAATGGCAATGCACTCATTTTGTATGAACTGCAGGATGATTTCTTA	1138
DB	1021	TCAGTGCAGGCTCAAAATGGCAATGCACTCATTTTGTATGAACTGCAGGATGATTTCTTA	1080
QY	1139	ATGAACCTTTAAAGATCTTAAATCCAGTCTGTATATGAGAGTGTTCACACTTCCAGTAAC	1198
DB	1081	ATGAACCTTTAAAGATCTTAAATCCAGTCTGTATATGAGAGTGTTCACACTTCCAGTAAC	1140
QY	1199	ATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTT	1258
DB	1141	ATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTT	1200
QY	1259	GGTCCATATGCCCCACAGGATGGACCCAACTATCAATGGGTGCTTATCAAGGAAGATC	1318
DB	1201	GGTCCATATGCCCCACAGGATGGACCCAACTATCAATGGGTGCTTATCAAGGAAGATC	1260
QY	1319	CCCTATCCAGGCGCAGGAACCTTGTCCAGCAAAACATTTGGTGGTTTTGACTCTACAAAG	1378
DB	1261	CCCTATCCAGGCGCAGGAACCTTGTCCAGCAAAACATTTGGTGGTTTTGACTCTACAAAG	1320
QY	1379	GACCTTCTGATGATTTTAAACCTTTGCAAGAACTCATCCAGCCATGTACAATCCAGTG	1438
DB	1321	GACCTTCTGATGATTTTAAACCTTTGCAAGAACTCATCCAGCCATGTACAATCCAGTG	1380
QY	1439	TTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAATTTCAATTTTACACAA	1498
DB	1381	TTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAATTTCAATTTTACACAA	1440
QY	1499	ATTGCTGTAGACCGAGTGGATGCAAGAGATGCAAGATGATGTTATGTTTATCGGAACA	1558
DB	1441	ATTGCTGTAGACCGAGTGGATGCAAGAGATGCAAGATGATGTTATGTTTATCGGAACA	1500
QY	1559	GATGTTGGGACCGTTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGA	1608
DB	1501	GATGTTGGGACCGTTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGA	1550

Search completed: July 31, 2003, 19:21:14

Job time : 7231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 13:13:31 ; Search time 326 Seconds
(without alignments)

4145.567 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aattctttatttatcatgatg.....aggetttttttcctaataacc 2709

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 424677 seqs, 249437934 residues

Total number of hits satisfying chosen parameters: 849354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	647.4	23.9	2347	1	PCT-US03-09929-95
2	647.4	23.9	2347	1	PCT-US03-09929-97
3	647.4	23.9	2373	1	PCT-US03-09929-93
4	647.4	23.9	6474	1	PCT-US03-09929-91
5	556	20.5	5231	6	US-10-357-930-23105
6	556	20.5	5231	6	US-10-357-930-28973
7	407.8	15.1	1671	6	US-10-357-820-27
8	126.6	4.7	1438	1	PCT-US03-09929-25
9	126.6	4.7	2113	1	PCT-US03-09929-55
10	125	4.6	1327	1	PCT-US03-09929-21
11	125	4.6	1492	1	PCT-US03-09929-23
12	125	4.6	1878	1	PCT-US03-09929-33
13	125	4.6	1821	1	PCT-US03-09929-15
14	125	4.6	1948	1	PCT-US03-09929-49
15	125	4.6	2583	1	PCT-US03-09929-51
16	125	4.6	2634	1	PCT-US03-09929-53
17	125	4.6	2995	1	PCT-US03-09929-19
18	125	4.6	3055	1	PCT-US03-09929-27
19	125	4.6	3106	1	PCT-US03-09929-17
20	125	4.6	3165	1	PCT-US03-09929-47
21	125	4.6	3498	1	PCT-US03-09929-31
22	125	4.6	4250	1	PCT-US03-09929-13
23	123.4	4.6	2944	1	PCT-US03-09929-29
24	112.4	4.1	1814	1	PCT-US03-09929-9
25	110.8	4.1	1914	1	PCT-US03-09929-1
26	110.8	4.1	1914	1	PCT-US03-09929-7

27	110.8	4.1	2049	1	PCT-US03-09929-3	Sequence 3, Appli
28	109.2	4.0	2558	6	US-10-369-072-15	Sequence 15, Appli
29	109.2	4.0	3112	1	PCT-US03-09929-5	Sequence 5, Appli
30	109.2	4.0	3112	6	US-10-369-072-17	Sequence 17, Appli
31	106	3.9	2739	1	PCT-US03-09929-11	Sequence 11, Appli
32	106	3.9	3868	6	US-10-369-072-13	Sequence 13, Appli
33	94.4	3.5	1908	1	PCT-US03-09929-35	Sequence 35, Appli
34	92.8	3.4	3983	1	PCT-US03-09929-45	Sequence 45, Appli
35	90.6	3.3	1923	6	US-10-461-791-12	Sequence 12, Appli
36	90.6	3.3	1923	6	US-10-461-792-12	Sequence 12, Appli
37	90.6	3.3	3261	6	US-10-461-791-5	Sequence 5, Appli
38	90.6	3.3	3261	6	US-10-461-792-5	Sequence 5, Appli
39	90.6	3.3	3694	6	US-10-461-791-3	Sequence 3, Appli
40	90.6	3.3	3694	6	US-10-461-792-3	Sequence 3, Appli
41	87.4	3.2	4234	6	US-10-451-010-12	Sequence 12, Appli
42	76.4	2.8	8062	6	US-10-273-573-3841	Sequence 3841, Ap
43	63.8	2.4	2271	6	US-10-369-072-7	Sequence 7, Appli
44	63.8	2.4	2281	6	US-10-369-072-9	Sequence 9, Appli
45	63.8	2.4	3042	6	US-10-293-244-3941	Sequence 3941, Ap

ALIGNMENTS

RESULT 1

PCT-US03-09929-95
; Sequence 95, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 95
; LENGTH: 2347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14)..(2338)
PCT-US03-09929-95

Query Match 23.9%; Score 647.4; DB 1; Length 2347;
Best Local Similarity 57.6%; Pred. No. 5.5e-198;
Matches 1265; Conservative 0; Mismatches 906; Indels 24; Gaps 5;

QY 218 ATTGCTGCTTTTCTGGGGAGTATTACTTACACAGACGAACTATCAGATGGGAG 277
DB 35 ATCACCTTGCTCCTGTGGGGTACTTACTGAGCTTTGGACAGAGGGTCATACAGCTGAT 94

Qy 278 AACAAATGTGCCAAGGCTGAATTAATCTTACAAAGAAATGTTGGAAATCCAAACAATGTGATC 337
Db 95 ACTACCACCCCGGTTACGCGTGTCAATAAAGAGCTCTTGAATCTGCAACAGAACATCA 154
Qy 338 ACTTTCATAGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCCTTTTGGATGAGAACGG 397
Db 155 ATAATTCATAGCCCTTTTGGATTTCTTGATCTCCATACAAATGCTGCTGGATGAATATCAA 214
Qy 398 AGTAGGCTGTATGTTGGAGCAAAAGGATCACATATTTTCAATCGACCTGGTTAAATATCAAG 457
Db 215 GAGAGGCTCTCGTGGGAGCGGACCTTGATATTTCCCTCAGCTTGGAGAGATCACT 274
Qy 458 GAT --- TTTCAAAGAATGTGTGCCAGTATCTTTACACGAAGAGAGATGAATGCAAGTGG 514
Db 275 GACGGCTATAAAGAGATACACTGGCCGAGTACAGCTCTAAAAAATGGAAGAAATGCATAATG 334
Qy 515 GCTGGAAAAGACATCTCGAAAGATGTCTTAATTTTCAATCAAGTACTTAAGGCATATAAT 574
Db 335 AAGGGAAAAGA --- TGGCGGTGAATGTGCAAAATTTATGTTCCGGTGTTCGATCACTATPAC 391
Qy 575 CAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCCAAATTTTCGACCTCATATTGAA 634
Db 392 AGGACACACCTCTTGA CCTGTGTGACTGTGAGCTTTTGATCCAGTTTGTGCTTCATCAGA 451
Qy 635 ATTGGAATCATCTCTGAGGACAAATATTTTTTAAGCTGGAGAACTCACATTTTGAACACGGC 694
Db 452 GTTGGATATCATTTGGAGGATCCCTCTGTTCACCTGGAATCACCCAGATCTGAGAGAGGA 511
Qy 695 CGTGGGAGAGTCCATATGACCCCTAAGCTGCTGCACAGCATCCCTTTTAAATAGATGGAGAA 754
Db 512 AGGGCAGATGTCTCTTTTGACCCAGCTCCTCTTCATCTCCACTTTAAATTTGGTAGTGA 571
Qy 755 TTATACTCTGGAACCTGCAGCTGATTTTATGGGGCGAGACTTTGTGTATCTTTCCGAACCTCT 814
Db 572 TTGTTTGTGTGACTCTACGTGACTACTGGACGAGAGACGCTGGGATCTTCGCGAGCATG 631
Qy 815 GGGCACCAACCAACCAATCAGSACAGACGAGCATGATTCAGAGTGGCTCAATGATCCAAAG 874
Db 632 GGGCGACTGGGCCATATCCGACCTGAGCATGACGAGCTGTGTTGAAAGAACCAAAA 691
Qy 875 TTCAATTAGTGCCCACTCATCTCAGAGAGTGACAACTCCTGAAGATGACAAAGTATACTTT 934
Db 692 TTTGTAGGTTCATACATGATTTCTCTGACAAATGAAGACAGAGATGACAAACAAGTATATTT 751
Qy 935 TTCTTCGCTGAAAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCAGCGCTAGAA 994
Db 752 TTTTCTACTGAAAGGCACCTGGAGGCAAAAAAATGCTCAGCAATTTACACAGGCTC 811
Qy 995 GGTGAGATATGCAAGAAATGACTTTGGAGGGCAGAGAGTCTGGTGAATAAATGGACAACA 1054
Db 812 GGGCGACTCTGTGTAATGATGTAGGAGGGCAGAGAAATCTGTTGTAATAAGTGGAGCACT 871
Qy 1055 TTCCCTCAAAGCTCGTCTGATTTGCTCAGTGGCAGGTCCAAATGGCATTTGACACTCAATTT 1114
Db 872 TTCCCTAAAGCGGAGACTCGTTTGTCTAGTACCAGGAATGAATGGAATTGACATATTTT 931
Qy 1115 GATGAACTGCAAGGATGATTTCTTAATGAACTTTTAAAGATCCTAAAAATCCAGTTGTATAT 1174
Db 932 GATGAATTAGAGACGTTTTTTTTTCTACCTACCGAGATCATAGAAATCAAGTGAATTT 991
Qy 1175 GGAGTGTTTACGACTCCAGTAACATTTTCAAGGGATCAGCCGCTGTGTATGTATAGCATG 1234
Db 992 GGACTCTTTAAACACTACCAAGTAAATTTTTTCGAGGGCATGCTATATGTGTCTATCAGATG 1051
Qy 1235 AGTGATGTGAGAGGGTGTCTTTGGTTCATATGTCACAGGAGATGGAACCAACTATCAA 1294
Db 1052 TCTAGCATTCGGGAGCGCTTTCAA CGGACCATATGTCATAGGAAGGACCTGTAATACCAC 1111
Qy 1295 TGGGTGCTTTTCAAAGGAAGGTCCTCTATCCAGGCCAGGAACCTTCTCCAGCAAAACA 1354
Db 1112 TGGTCAGTCTATGAAGGAAGAAGTCCCTTATCCAGGGCTCGTTTCTGTGTGCGCAGCAAGTA 1171
Qy 1355 TTTTGGTG --- GTTTTGTACTCTACAAAGGAGCCTTCTCGATGATGTATTAACCTTTGCAAGA 1411

DB	1172	AATGGAGGGAGATACGGAAACCAACAGGACTATCTCTGATGATGCCATCCGATTTGGCAAGA	1231
QY	1412	AGTCATCCAGCCATGTACAAATCAGTGTTCCTATGAACAAATCGCCCAATAGTGATCAAA	1471
DB	1232	AGTCATCCACTAATGTACAGGCCATAAACCCTGCCCATATAAAACCAATATTAGTAAAA	1291
QY	1472	ACGGATGTAAATATTCAATTTTACACAAATTTGCTAGACCGAGTGGATCGAGAAGATGGA	1531
DB	1292	ACAGTGGAAAAATATTAACCTGAAACAAATAGCAGTAGATCGAGTGGAAAGCTGAGGATGGC	1351
QY	1532	CAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCCTTAAAGTAGTTTCAATT	1591
DB	1352	CAATATGACGTCCTGTTTATTGGGACAGATAATGGAAATTTGCTGTAAGAGTAATCACAATT	1411
QY	1592	CCTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCTGCTGGAAGAAATGACAGTTTTTTCG	1651
DB	1412	TACAAACCAAGAAATGGAAATCAATAGGAAGAAAGTAAATCTTAGAAGAACTTCAGATATTC	1471
QY	1652	GAACCGACTGCTATTTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACCACTATATATGGT	1711
DB	1472	GATCCAGTTCCTATTATTTCTATGGAGATTCTTCAAAGCGGCAACAGCTGTATATTGGA	1531
QY	1712	TCAAAGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTACCGGAAGCGGTCT	1771
DB	1532	TCTGCTTCTGCTGGCTCAAGTCAGATCCATCACTGTGAATGTATGGAAGTGCTTGT	1591
QY	1772	GCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTTGGATGGTTCCTCATGTTCTCGC	1831
DB	1592	GCTGACTGCTGCTCGGCTCGAGACCCCTTACTGTGCTGGGATGGCATATCTTGCTCCCG	1651
QY	1832	TATTTTCCCAAC-----TGCAAAAGAGACGCAAGACGCAAGATATAAGAAATGGA	1882
DB	1652	TATTTACCAACAGGCAACACATGCAAAAGCGTTTCCGGAGACAAGATCTCGCATGGA	1711
QY	1883	GACCCACTGACTCACTGTTTCAGACTTACACCATGATATCAACATGGCCACAGCCCTGAA	1942
DB	1712	AATGCACTCAGCAGTGCCTTTGGACAACTGTTGTTGGGGATGCTTTGGATGAAGACTGAA	1771
QY	1943	GAGAGAATCATATGCTGTAGAGAATAGTAGCACATTTTTCGAAATGCACTCCGAAGTCG	2002
DB	1772	GAACATCTGGCTATTGGCATAGAAACAACAGTACTTTGCTGGAAATGATACCCACGATCT	1831
QY	2003	CAGAGACGCTGCTCTATTGGCAATTTCCAGAGCGCAAAATGGAAGCGCAAAAGAGAGATC	2062
DB	1832	TTCAAGCGAAAGTTATCTGTTTGTACAGAAAGGACGTGAGACAAGAAAGAGGAGGTG	1891
QY	2063	AGAGTGGATGATCATATCATCAGGAACAGATCAAGGCTTCTGCTAGTAGTCTACAAACG	2122
DB	1892	AAGACAGATGACAGAGTGGTTAAGATGGACCTTGGTTTACTCTTCTTAAGGTTTACACAA	1951
QY	2123	AAGGATTTCAAGCAATTTACCTCTGCCATCGGCTGGAAACATGGGTTTCATACAAACTCTTCT	2182
DB	1952	TCAGATGCTGGGACCTATTTTGGCCAGACAGTAGAGCATAGCTTTGTCCATACGGTCCGT	2011
QY	2183	AAGGTAAACCTCGGAATGTCATACACAGAGCATTTGGAAAGAACTTCTTCATTAAGAGATGAT	2242
DB	2012	AAAATCACCTTGGAGGTAGTGGAAAGGAGAAAGTCGAGGATATGTTTAAACAAGACGAT	2071
QY	2243	GATGGAGATGGCTCTTAAGACCAAGAAATGTC-----CAATAGCATGACACCTAGCCAG	2296
DB	2072	GAGGAGACAGGCGATCACAGGATGCGCTGCTCTCAGAGTAGCATCTCGCAGGAGCA	2131
QY	2297	AAGGTCTGGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGATGGAT	2356
DB	2132	AAACCATGGTACAAGGAATTTCTTGACGCTGATCGGTTATAGCACTTCCAGAGAGTGGAA	2191
QY	2357	GAGTTCGTGAAACAAGTTTGGAAAAAGGGACCGAAA	2391
DB	2192	GAATATCGCAGAAAGTATGTTGCAAGATAGAAA	2226

RESULT 2

Query Match	23.9%;	Score 647.4;	DB 1;	Length 2347;
Best Local Similarity	57.6%;	Pred. No. 5.5e-198;		
Matches 1265;	Conservative	0;	Mismatches 96;	Indels 24; Gaps 5
QY	218	ATTGTCTGTCTTTCTTGGGGAGTATTACTTACAGCAAGACAAACTATCAGAAATGGGAG	277	
DB	35	ATCACCTTGCCTCTGTGGGGTTACTTACTGGAGCTTTGGACAGAGGTATACAGCTGAT	94	
QY	278	AACATGTGCCAAGCTGAAATTATCTTACAAGAAATGTTGGAATCCAACAAATGTGATC	337	
DB	95	ACTACCCACCCCGGTTACGCGTCTGCATAAAGAGCTCTTGAATCTGAACAGAACATCA	154	
QY	338	ACTTTCAATGGCTTGGCCAAACAGCTCCAGTTTATCATACCTTCCTTTGGATGAGNACGG	397	
DB	155	ATATTTCATAGCCCTTTTGGATTTCTTGATCTCTACAATGCTGCTGGATGAATATCAA	214	
QY	398	AGTAGCTGTATGTTGGACAAAGATCACATATTTTTCATTGCACCTGGTTAATATCAAG	457	
DB	215	GAGAGGCTCTTGGTGGAGGCAGGAGCCTTGATATATTCCTCAGCTTGGAGAGAAATCA	274	
QY	458	GAT---TTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAAATGCAATGG	514	
DB	275	GACGCTATAAGAGATACACTGGCCGAGTACAGCTCTAAAATGGAAGAATGCATATG	334	
QY	515	GCTGAAAAGACATCCTGAAAAGATGTGCTAAATTCATCAAGGTACTTAAGGCATATAAT	574	
DB	335	AAGGGAAAAGA---TTCGGGTGAATGTGCAAAATATGTTCCGGGTTTGGCATCACTATAAC	391	
QY	575	CAGACTCATCTGTACCGCTGTGGACGGGGCTTTTCATCCAAATTTGCACCTACATTGNA	634	
DB	392	AGACACACCTTCTGACCTGTGGTACTGGAGCTTTTGTATCCAGTTTGTGCGCTTCATCA	451	
QY	635	ATTGGACATCATCTCCCTGAGGACAAATATTTTTAAAGCTGGAGAACTCACATTTTGAAACCGC	694	


```
Db 1532 TCTGCTTCTGCTGGCTCAAGTCAGATTCATCACTGTGCATGTATGGAAGTCTTGT 1591
Qy 1772 GCTGAGTGTGCTCGCCGAGACCTTACTGTGCTTGGATGGTTCATGCTTCTCGC 1831
Db 1592 GCTGAGTGTGCTCGCCGAGACCTTACTGTGCTTGGATGGTTCATGCTTCTCGC 1651
Qy 1832 TATTTTCCAC-----TGCAAGAGACGCAAGACGCAAGATATAAGAAATGGA 1882
Db 1652 TATTACCAACAGGCACATGCAAAAGCGTTTCCGGAGACAAGTGTTCGACATGA 1711
Qy 1883 GACCCACTGACTCATCTGTTACACTTACCATGATAATCAATGACGACCCCTGAA 1942
Db 1712 AATGCAGCTCAGCAGTGTCTTGACACACAGTTTGTGGGATGCTTTGATGAAGTGA 1771
Qy 1943 GAGAGAATCATCTATGCTGTAGAGATAGTAGACATTTTGGATGCACTCGAGTCG 2002
Db 1772 GAACATCTGCTTATGGCATAGAGAACACAGTACTTGTGGAATGTACCCACAGATC 1831
Qy 2003 CAGAGAGCGTGTCTATTGGCAATTCAGAGCGCAAAATCAAGAGCGAAAGAGATC 2062
Db 1832 TTACAAGCGAAGTATCTGTTTGTACAGAAAGGACGTGAGACAAAGAGGAGTG 1891
Qy 2063 AGAGTGGATGATCATATCATACGAGACAGATCAAGGCTTCTGCTAGTGTCTACAAC 2122
Db 1892 AAGACAGATCAGAGAGTGGTTRAGATGACCTTGGTTTACTCTTCTTAAGGTTACAAA 1951
Qy 2123 AAGGATTCAGGCAATTAACCTCTGCCATGCTGGTGGACATGGTTCATAAACTCTT 2182
Db 1952 TCAGATGCTGGGACCTATTTTGGCCAGACAGTAGATAGCTTTGTCTCATACGGTCC 2011
Qy 2183 AAGGTAACCTGGAAGTCTATGACACAGACAGCAATTTGGAAGAACTCTTCTATAAGATGAT 2242
Db 2012 AAAATCACTTGGAGTGTGGAGAGGAGAAAGTCGAGATATGTTTAAACAGACGAT 2071
Qy 2243 GATGGAGATGGCTCTTAAGACCAAGAAATGTC-----CAATAGCATGACACTAGCCAG 2296
Db 2072 GAGGAGGACAGGCATCACAGGATGCTTGTCTGCTCAGAGTAGCATCTCGCAGGAGCA 2131
Qy 2297 AGGTCGTGTACAGACATTCATGACGCTCATCAACCCCACTCAACGATGAT 2356
Db 2132 AAACCATGTTACAGGAATCTTTCGACGTGATCGGTTATAGCAACTTCCAGAGTGGAA 2191
Qy 2357 GAGTCTGTGCAACAGTTTGGAAAAGGACCGAAA 2391
Db 2192 GAATCTGCGAGAAAGTATGGTGCACAGATAGAA 2226

RESULT 3
PCT-US03-09929-93
; Sequence 93, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
```

```
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 93
; LENGTH: 2373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)...(2347)
PCT-US03-09929-93

Query Match 23.9%; Score 647.4; DB 1; Length 2373;
Best Local Similarity 57.6%; Pred. No. 5.5e-198;
Matches 1265; Conservative 0; Mismatches 906; Indels 24; Gaps 5;

Qy 218 ATTTGCTGCTTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAG 277
Db 44 ATCCTTGTCTTCTGTGGGGTTACTTACTGGAGCTTTGGACAGGAGGTCTACAGCTGAT 103.
Qy 278 AACAAATGTGCCAAGGCTGAAATTTATCTTACAAAGAAATTTTGGAAATCCAAATGTGATC 337
Db 104 ACTACCCACCCCGGTACGCTGTCAATAAAGAGCTCTTGAATCTGAACAGAACATCA 163
Qy 338 ACTTTCAATGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTTTGGATGAGAAACGG 397
Db 164 ATATTTATATGCTCTTTTGGATTTCTTGAATCTTCAATCAATGCTGCTGGATGAATATCAA 223
Qy 398 AGTAGGCTGTATTTGGAGCAAGGATCACATATTTTCAATTCGACCTGTTAAATATCAAG 457
Db 224 GAGAGGCTCTTCTGTGGAGGAGGAGCCTTGATATTTCCCTCAGCTTGGAGAAATCAGT 283
Qy 458 GAT---TTTCAAAAGATTGTGTGGCCAGTATCTTTACACAGAGAGATGAATGCAAGTGG 514
Db 284 GACGCTATAAAGAGATACACTGGCCGAGTACAGCTCTAAAAATGGAAGAAATGCATAATG 343
Qy 515 CTGGAAGACATCTCGAAAGATGTCTTAATTTCAATCAAGGTACTTAAAGCATATATAT 574
Db 344 AAGGGAAAGA---TGCAGGTTGAATGTGCAAAATATATGTTTCGGGTTTTCATCATATAAC 400
Qy 575 CAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCAATCCAAATTTGACCTACATTTGAA 634
Db 401 AGACACACCTTCTGACCTGTGGTACTGGAGCTTTTGAATCCAGTTTGGCTTCTATCAGA 460
Qy 635 ATTGACATCATCTCGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAACGGC 694
Db 461 GTTGATATCATTTGGAGGATCCTCTGTTTCACTGGAATCACCCAGATCTGAGAGAGGA 520
Qy 695 CGTGGAAAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGAA 754
Db 521 AGGGGCGATGCTCCTTTTGGACCCAGCTCTCTTCACTTCCACTTTTAAATTTGGTAGTGA 580
Qy 755 TTATCTCTGGAACCTGCGAGCTGATTTTATGGGCGAGACTTTTGTCTATCTTCCGAACTCT 814
Db 581 TTGTTGCTGCACTCTACAGTGACTACTGGAGAGAGACGCTGGATCTTCCGAGCAATG 640
Qy 815 GGGCACCACCAACCAATCAGGACAGAGCATGATTCCAGGTGGCTCAATGATCCAAAG 874
Db 641 GGGGACCTGGCCCATATCCGCACTGAGCATGACCATGAGCGTCTGTTGAAAGAACCAAAA 700
Qy 875 TTCAATAGTCCCACTCTCATCTCAGAGATGACAACTCTGAAAGTACAAAGTATATCTTT 934
Db 701 TTTGTAGGTTTCATCATCATGATTTCTGCAATGAAGACAGAGATGACAAAGATATATTTTC 760
Qy 935 TTCTTTCCGTGAAATGCAATAGATGGAGAACACTCTCGAAAGCTACTCACGCTAGATA 994
Db 761 TTTTCTACTGAAAGGCACTGGAGGAGAAACAAATGCTCAGCAATTTACCCAGGGTTC 820
Qy 995 GGTGAGATATGCAAGAAATGACTTTTGGAGGGGACAGAAAGTCTGTGGTGAATTAATGGAACA 1054
```

Db 821 GGGCGACTCTGTGAATGATGTAGGAGGCGCAGAGAACTACTGGTGAATTAAGTGGAGCACT 880
Qy 1055 TTCTTCAAAAGCTGGTCTGATTTGCTGAGTCCAGTCCAAATGGCAATGACATCACTATTT 1114
Db 881 TTCTTAAAGCGAGACTCGTTTGCTCAGTACCAGGAATGAATGGAATTCACACATATTTT 940
Qy 1115 GATGAAGTCCAGAGTATTTCTTAATGAATTTAAAGATCCCTAAATCCAGTGTATAT 1174
Db 941 GATGAATTAAGAGAGCTTTTGTGTACCTACAGAGATCAATGAATCCAGTGTATTT 1000
Qy 1175 GGAGTGTTTACCACTTCCAGTAACATTTTCAAGGATCAGCGTGTGTATGTATAGCATG 1234
Db 1001 GGCATCTTTAACTACCACTACCACTATTTTTCAGGCGATGCTATATGTCTATCACATG 1060
Qy 1235 AGTGAATGAGAGGCTGCTTCCATATGCCCAGGAGTGGACCCCACTATCAA 1294
Db 1061 TCTAGCATTCGGGCGACCTTCAACGGACCATATGCACATAAGGAAGGACCTGAATACCAC 1120
Qy 1295 TGGTGCCTTTCAAGGAAGTCCCTATCCAGGCGCAGGAACCTTGTCCAGGAAACA 1354
Db 1121 TGGTCAGTCTATGAAGGAAGTCCCTTATCCAGGCGCTGTTCTTGTCCAGCAAGTA 1180
Qy 1355 TTGTGTG---GTTTGTACTCTACAAAGGACCTTCTGTGATGATGTTATAACCTTTGCAAGA 1411
Db 1181 ATGGAGGAGATACGGAACACACAGGACTATCTGTATGATGCCATCGATTCGATCAAGA 1240
Qy 1412 AGTCATCCAGCATGTACAAATCCAGTGTTCCTATCAACAATCGCCCAATAGTATCAAA 1471
Db 1241 AGTCATCCACTAATGTACAGGCCATAAAACCTGCCCCATAAAACCAATATTAGTAAA 1300
Qy 1472 ACGGATGTAATATCAATTTACAAATTTGCTGAGACCGAGTGGATGCAGAAATGGA 1531
Db 1301 ACAGATGGAATAATCAACTGAACCAATAAGCAGTAGATCGAGTGGAGCTGAGATGCG 1360
Qy 1532 CAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTTAAAGTACTTTCAATT 1591
Db 1361 CAATATGACGCTGTTTATGCGACAGATAATGGAATGTGCTGAAAGTATCACAATT 1420
Qy 1592 CTAAGGAGACTTGTGTATGATTTAGAAGAGTTCCTGCTGGAAGAAATGACAGTTCGCG 1651
Db 1421 TACAACCAAGAAATGGAATCAATGGAAGAAATTTCTAGAAGAACTTCAGATATTCAG 1480
Qy 1652 GAACGACTGTATTTTACGAATGAGCTTCCACTAAGCAGCAACAATATATATTTGGT 1711
Db 1481 GATCAGTTCCTATATTTCTATGAGATTTCTTCAAAAGCGCAACAGCTGTATATGGA 1540
Qy 1712 TCAACGGCTGGGTTGCGCCAGCTCCCTTTACACCGGTGTGATTTACGGGAAGCGGT 1771
Db 1541 TCTGCTTCTGCTGGCTCAAGTCAAGATTCATCACTGTGACATGTATGGAAGTCTTGT 1600
Qy 1772 GCTGAGTGTGCTCGCCGACACCTTACTGTGCTGGGATGGTTCCTGCATGTTCTCGC 1831
Db 1601 GCTGACTGCTGCTGCTCGACACCTTACTGTGCTGGATGGCATATCTGCTCCCGG 1660
Qy 1832 TATTTTCCAC-----TGCAAAGACGCAACAGACGCAAGATATAGAAATGGA 1882
Db 1661 TATTACCCCAACAGGCACATGCAAAAAGCGCTTTCCGAGACAAAGATGTTCCGACATGGA 1720
Qy 1883 GACCCACTGACTCTGCTGACTTACACCATGATATCACTCCATGGCCAGCCCTGAA 1942
Db 1721 ATGCAGCTCAGCAGTCTTTGGAACAACAGTTTGTGGGATGCTTTGGATAAGATGAA 1780
Qy 1943 GAGAGATCATCTATGCTGTAGAGAAATAGTACACATTTTGGAAATGAGTCCGAAGTCG 2002
Db 1781 GAACATCTGGCTTATGGCATAGAAACACAGTACTTTGCTGGATGTACCCGACGATCT 1840
Qy 2003 CAGAGAGCGCTGCTTATTTGGCAATTCAGAGCGGAAATGAAGCGCAAAAAGAGAGATC 2062
Db 1841 TTACAAGCGAAAGTATCTGTTTCTGACAGAAAGGACGCTGACAGAAAGAGAGAGGCTG 1900
Qy 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGSCCTTCTGCTACGTAGTCTACAAAG 2122
Db 1901 AAGACAGATGACAGAGTGGTAAAGTGGACCTTGTGTTTACTTCTTCCTAAGGTTACACAA 1960

Qy 2123 AAGANTTCAGCAATATACCTCTGCCATCGGTGGAAATGGTTTCATACAACTCTTCTT 2182
Db 1961 TCAGATGCTGGGACCTATTTTTCAGACAGCTAGAGCATAGCTTTGTCCATACGCTCCGT 2020
Qy 2183 AAGTTAACCTTGGAGTCTATTCACACAGCATTTGGGAAGAACTTCTTCATAAAGATGAT 2242
Db 2021 AAAATCACCTTTGGAGGTAGTGGAAAGAGGAAAGTTCGAGGATATATTTTAAACAAGACCGAT 2080
Qy 2243 GATGAGATGCTCTTAAGACCAAAAGAAATGTC-----CAATAGCATGACACCTAGCCAG 2296
Db 2081 GAGGAGCAGCGCATCACAGATGCCTTGTCTGCTCAGATGAGCATCTCGCAGGAGCA 2140
Qy 2297 AAGTCTGTGACAGACTTTCATGAGCTCATCAACCCCAATCTCAACACAGATGAT 2356
Db 2141 AAACCATGTTCAAGGAATTTCTGAGCTGATCGTTTATAGCAACTTCCAGAGAGTGGAA 2200
Qy 2357 GAGTCTCTGAACAAGTTTGGAAAAGGACCGAAA 2391
Db 2201 GAATACTCGGAGAAAGTATGGTGCACAGATAGAAA 2235

RESULT 4

PCT-US03-09929-91
; Sequence 91, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSequid version 0.1
; SEQ ID NO 91
; LENGTH: 6474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (467)..(2791)
PCT-US03-09929-91

Query Match 23.9%; Score 647.4; DB 1; Length 6474;
Best Local Similarity 57.6%; Pred. No. 9.2e-198;
Matches 1265; Conservative 0; Mismatches 906; Indels 24; Gaps 5;

Qy 218 ATTCTGTCTCTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAAAATGGGAAG 277
Db 488 ATCCTTGTCTCTCTGGGTTACTTACTGGAGTTTGGACAGGAGGTCTACAGCTGAT 547
Qy 278 AACAAATGCGCAAGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAACAAATGTGATC 337

Db 548 ACTACCCACCCCGGTTAGCCCTGTGCATATAAGAGCTCTTGAATCTGAACAGACATCA 607
Qy 338 ACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGNACGG 397
Db 608 ATATTTTCATAGCCCTTTTGGATTTCTTGATCTCCATACAAATGCTGCTGGATGAATCA 667
Qy 398 AGTAGCTGTATGTTGGAGCAAGGATCAATATTTTCATTCGACCTGCTGTTAATATCAAG 457
Db 668 GAGAGGCTCTTCTGGAGGCGAGGACCTTGATATTCCTCAGCTTGGAGAGATCAGT 727
Qy 458 GAT---TTTCAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514
Db 728 GACGGCTATAAAGAGATACACTGGCCGAGTACAGCTCTAATAATGGAAGATGCAATAG 787
Qy 515 GCTGGAAGAGACATCTGGAAGATGCTGCTAATTTTCATCAAGTACTTTAAGCCATATAT 574
Db 788 AAGGGAAGA---TGCGGCTGAATGTGCAATATATGTTCCGGTTTTGCATCACTATAC 844
Qy 575 CAGACTCACTTGTACGCTGTGGAACGGGGCTTTTTCATCCAAATTTGCACCTACATTTGAA 634
Db 845 AGGACACACTTCTGACCTGTGTACTGGAGCTTTTGATCCAGTTTGTGCTTTCATCAGA 904
Qy 635 ATTGGACATCATCTGAGGACAAATATTTTAAAGTGGAGAACTCAATTTTGAACACGGC 694
Db 905 GTTGGATATCATTTGGAGGATCTCTGTTTCACTTGAATCACTGGAATCACTGAGAGAGA 964
Qy 695 CGTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCTCTTTTAAATAGAGAGAA 754
Db 965 AGGGGAGATGCTCTTTGACCCAGCTCTCTCTTCATCTCCACTTTAATTTGGTAGTGA 1024
Qy 755 TTATACTCTGGAAGTGCAGCTGATTTTATGGGCGAGACTTTTGTCTATCTTCCGAACTCTT 814
Db 1025 TTGTTGCTGGACTCTACAGTACTACTGAGCAGAGAGCTGCGATCTTCCGAGCATG 1084
Qy 815 GGGACCAACCACTTACAGAGAGAGAGAGATGATTCAGGTGCTCAATGATCAAAG 874
Db 1085 GGGCGACTGGCCCATATCGCACTGAGCATGACGATGAGCGTCTGTTGAAAGAACAAA 1144
Qy 875 TTCAATAGTCCACCTCATCTCAGAGTGTCAATCTCGAAGATGACAAAGTATACTTT 934
Db 1145 TTGTAGGTTTACATGATTTCTGACATGAAGAGAGAGATGACAAAGATATATTTTC 1204
Qy 935 TTCTTCGTTGAAATGCAATAGATGAGAACACTCTGGAAGAGTACTCACCGCTAGAATA 994
Db 1205 TTTTITACTGAGAGGCACTGGAGCAGAGAAATGCTCAGCAATTTACACAGGGTC 1264
Qy 995 GGTGAGATGCAAGATGACTTTGGAGGCGACAGAGTCTGGTGAATAAATGGAACA 1054
Db 1265 GGGCGACTCTGTGTAATGATGTAGGAGGCGAGAGAACTACTGGTGAATAAGTGGAGCA 1324
Qy 1055 TTCTCAGAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGCAATGACACTCATTTT 1114
Db 1325 TTCTTAAAGCGAGACTCGTTGCTCAGTACCAGGAATGAATGGAATTTGACACATATTT 1384
Qy 1115 GATGAATCTCAGGATGATTTCTCAATGAATTTTAAAGATCTTAAAGATCTTCAATTTAT 1174
Db 1385 GATGAATTAGAGGACGTTTTTTTGTCTACCTACAGAGATCAATGAATCCAGTATTT 1444
Qy 1175 GGAGTGTTCAGACTTCCAGTAACTTTTCAAGGAGATCAGCGTGTGTATGATAGCATG 1234
Db 1445 GGACTCTTAAACACTACCAGTAAATTTTTCGAGGCGATGCTATATGTTCTATCACATG 1504
Qy 1235 AGTGATGTGAGAGGTTCTTGTGCTATATGCCCAGGGATGGACCCCACTATCA 1294
Db 1505 TCTAGCATTCGGGCGCTTCAACGGACCATATGACATAAGGAAGGACCTGAAATACCAC 1564
Qy 1295 TGGGTGCTTTTCAAGGAAGAGTCCCTTATCCACGGCAGGAACTTTGTCACCAAAACA 1354
Db 1565 TGGTCAGTCTATGAAGGAAGTCCCTTATCCAGGCTGGTTCTTGTGTCAGCAAGATA 1624
Qy 1355 TTGTGTG---GTTTGTACTCAAGAGGACCTTCTGTGATGATGTTTAACTTTGCAAGA 1411
Db 1625 AATGGAGGAGATACGGAAACCAAGGACTATCTCTGATGATGCTCATCGGATTTGCAAGA 1684

Qy 1412 AGTCATCCAGCATGTACAATCCAGTGTTCCTATGAACAATCGCCCAATAGTGTATCAAA 1471
Db 1685 AGTCATCCATTAATGTACAGGCCATAAACTGCCCATAAACCAATATTTGGTAAAA 1744
Qy 1472 ACGGATGPAATATTAATCAATTTTACAAATTTCTGTAGACCGAGTGGATGCAAGATGA 1531
Db 1745 ACAGATGGAATAATATAACCTGAAACNAATAGCAGTAGATCGATGGAAGCTGAGGATGC 1804
Qy 1532 CAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCCTTAAAGTAGTTCATTT 1591
Db 1805 CAATATGACGCTCTGTTTATTTGGGACAGATAATGGAATTTGCTGAAAGTAAATCAATTT 1864
Qy 1592 CTTAAGGAGACTTGTGTATGTTTGAAGAGGTTCTGCTGGAAGAATGACAGTTCCTGG 1651
Db 1865 TACAACCCAGAAATGGAATCAATGGAAGAATTAATCTAGAGAATCTCAGATATTCAG 1924
Qy 1652 GAACCGACTGCTATTTTACGAATGGAGCTTTCCATTAAGCAGCAACAATATATATGGT 1711
Db 1925 GATCCAGTTCTATTTATTTCTATGGAGATTTCTTCAAACGGCAACAGCTGTATATTGA 1984
Qy 1712 TCAAGGCTGGGTTGCCAGCTCCCTTTACACGGTGTGATATTTACGGGAACCGTGT 1771
Db 1985 TCTGCTTCTGCTGGCTCAAGTCAGATTCATCACTGTGACATGTATGGAAGTCTTGT 2044
Qy 1772 GCTGAGTGTCCCTCGCCGAGAGCCCTTACTGTGCTTGGGATGGTTCTGCAATGTTCTCG 1831
Db 2045 GCTGACTGCTGCTGCTGAGACCTTACTGTGCTGGGATGGCATATCTGCTCCCG 2104
Qy 1832 TATTTTCCAC-----TGCAAGAGAGCGCAAGACAGCAAGATATTAAGAAATGA 1882
Db 2105 TATTACCCAAACAGGACACACATGCAAAAAGGCGTTTCCGAGACAAGATGTTCCGACATGA 2164
Qy 1883 GACCCACTGACTACTGTTTACAGCTTACACCATGATTAATCACCATGGGCCACAGCCCTGA 1942
Db 2165 AATGAGCTCAGCAGTGTCTTTGGAACAAGTGTGTTGGGATGCTTTGGATAAGACTGAA 2224
Qy 1943 GAGAGATCACTATGTTGTAGAGAAATAGTAGACATTTTGGAAATGAGTCCGAAGTCG 2002
Db 2225 GAACTCTGGCTTATGGCATAGAGAACAACTACTTTGCTGGAAATGTACCCACGATCT 2284
Qy 2003 CAGAGAGCGCTGCTTATTTGGCAATCCAGAGGCGAAATGAAGAGCGGAAAAAGAGATC 2062
Db 2285 TTACAAGCGAAAGTTATCTGCTTTGTACAGAAAGGACGTCAGACAAGAAAAAGAGAGGTG 2344
Qy 2063 AGAGTGAATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAGTCTACAAAG 2122
Db 2345 AAGACAGATGACAGAGTGGTTAAGATGGAACCTTGGTTTACTCTTCTTAAGGTTACACAA 2404
Qy 2123 AAGGATTCAGGCAATTTACTCTGCTGCGGTGGAACATGGGTTTCATACAAATCTTCTT 2182
Db 2405 TCAGATGCTGGACCTATTTTTCAGACAGTAGTAGCATAGCTTTTGTCCATACGCTCGT 2464
Qy 2183 AAGGTAAACCTTGGAGTCAATTCACACAGAGCATTTTGAAGAACTTCTTCAATAAGATGAT 2242
Db 2465 AAAATCACTTTGGAGGTAGTGAAGAGAGAGAAAGTCGAGGATATGTTTAAACAAGGACAT 2524
Qy 2243 GATGAGATGCTCTAAGACCAAGAAATGTC-----CAATAGCATGACCTAGCCAG 2296
Db 2525 GAGGAGGACAGGCAATCAGGATCCCTTGTCTCTCTCAGAGTAGCATCTCCGAGGAGCA 2584
Qy 2297 AAGGTCTGGTACAGAGACTTTCATGAGCTCATCAACCAACCCCAATCTCAACACGATGAT 2356
Db 2585 AAACCATGTTACAGGAATTTCTTCCAGCTGATCGTTATAGCAACTTCCAGAGTGA 2644
Qy 2357 GAGTCTGTGAACAAAGTTTGGAAAAGGACCGAAA 2391
Db 2645 GAACTACTGCGAGAAAGTATGTTGTCACAGATAGAAA 2679

RESULT 5
US-10-357-930-23105
; Sequence 23105, Application US/10357930

GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23105
; LENGTH: 5231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 5231
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23105

Query Match 20.5%; Score 556; DB 6; Length 5231;
Best Local Similarity 57.6%; Pred. No. 2.5e-168;
Matches 1117; Conservative 0; Mismatches 800; Indels 21; Gaps 6;

QY	288	CAAGCTGAAATATCTACAAAGAAATGTTGGAATCCAAATGATGATCATTCAATG	347
Db	642	CAAGAGTTTATTAATTTGATGAACTTCGAGAAACCAAGACCTCTGAAATCTCAGCC	701
QY	348	GCTTGGCCCAACAGCTCCAGTTATACATCTCTTGGATGAGGACGGAGTAGGCTGT	407
Db	702	TTTCCCAACATCTCTTAGACTACAGGATTTATTAATGATGAAATCAGGACCGAT	761
QY	408	ATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAAATCA---AGGATTTTC	464
Db	762	ATGTTGGGAAGCAAGATCACATTTCTCCCTGAAATATTAACAAATATAAGTCAAGAAGCTT	821
QY	465	AAAAGATTGTGGCCAGTATCTTACACAGAGAGATCAATGCAAGTGGGCTGGAAG	524
Db	822	TGAGTGTCTTCTGGCCAGATCTCAATCAAAAGTTGAAGATGCAAAATGGCTGGCAAG	881
QY	525	ACATCTCTGAAAGATGTGCTAATTTTCATCAAGGTACTTAAGGCATATATCAAGCTCACT	584
Db	882	ATCCCAACACAGCTGTGGAACTTTGTCGGTGAATTCAGACTTTCATCGCACATTT	941
QY	585	TGTACGCTGTGGAAACGGGGCTTTTCATTCCAATTTGCACTTACATTAATGAAATGGAATC	644
Db	942	TGTATGTCTGTGGAGTGGCGCTTTTCAGTCTCTGTCTGTACTTACTTGAACAGAGGAGGA	1001
QY	645	ATCCTGAGGACAAATTTTAAAGTGGAGATCAATTTTGAACCGCGTGGGAAGA	704
Db	1002	GATCAGAGACCAAGTTTTCATGATG---ACTCCAAGTGTGAATCTGGAAGAGGACGCT	1058
QY	705	GTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAATAGATGGAAATATATCTCTG	764
Db	1059	GCTCTTTCACCCCAAGTGACAGGCTGTCTGTTATGATCATCATGAGGAGCTTTCTCTG	1118
QY	765	GAACTGACGTGATTTTATGGGGGAGAGCTTTTGCTATCTTCGAACTCTTGGGCCACC	824

Db	1119	GAATGTATATAGATTTTCATGGGACAGATGCTGCTATTTTTCGAAGTTTAAACCAAGAGA	1178
QY	825	ACCAATCAGACAGACAGCATGATTCAGCTGGCTCAATGATCCAAAGTTCATTAGTG	884
Db	1179	ATGCGGTGAGAACTGATCAACAATATTCCAATGGCTAAGTGAACTATGTTGTAGATG	1238
QY	885	CCCACCTCATCTCAGAGAGTGACAAATCCTGAAGATGACAAAGATATCTTTTCTCCGTG	944
Db	1239	CACATGTCTATCCAGATGGTACTGATCCAAATGATGCTAAGGTGATCTTCTTCTCAAAG	1298
QY	945	AAATGCAATAGATGAGAGACACTCTGGAAGCTTACTCAGCTAGATAGTGCAGATAT	1004
Db	1299	AAAACTGACTGACAAATAACAGAGACGAAACAGATTCATTCCATGATTTGCTCGAATAT	1358
QY	1005	GCAAGATGACTTTGGAGGGACAGAACTCTGGTGAATAAATGGAACAACATTCCTCAAAG	1064
Db	1359	GTCTTAATGACACTGGTGGACTGCGTAGCTTGTCAACAGTGGACCACTTCTTAAGG	1418
QY	1065	CTCGTCTGATTTGCTCAGTGCAGGTCCAAATGCAATTTGACACTCATTTTGAAGACTGC	1124
Db	1419	CGAGCTGGTGTCTCGGTAAACAGATGAAGACGGCCAGAAACACACTTTGATGAATTAG	1478
QY	1125	AGATGTATTCCTTAATGAACTTTAAAGATCTTAAATCCAAATCCAGTTGTATATGAGGT	1184
Db	1479	AGGATGTGTTTCTGCTGGAAACTGATAACCCGAGGACCAACACTAGTGTATGGCATTTTA	1538
QY	1185	CGACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGATGATGTA	1244
Db	1539	CAACATCAAGCTCAGTTTCAAGATCAGCGGTGTGTGTATCAATTTCTGATATAC	1598
QY	1245	GAAGGTGTCTCTTGGTCCATATGCCACAGGGATGGACCAACTATCAATGGGTGCCTT	1304
Db	1599	AGACTGTGTTTAAATGGGCTTTTGGCCACAAAGAGGGCCCAATCATCAGCTGATTTCT	1658
QY	1305	ATCAAGGAGAGTCCCTTATCCACGGCCAGAACTTGTGCCAGCAAAACATTT---GGTG	1361
Db	1659	ATCAGGCGAAGTTCCATATCTCGCCCTGGAACTTGTCCAGGAGGACATTTACACCCA	1718
QY	1362	GTTTGTACTPACAAAGGACCTCTCTGATGATGTTTATAACCTTTTCAAGAGATCATCCAG	1421
Db	1719	ATATCGGNACACAGAGGATTCAGATGATGTTGTCACTTTTATTCGGAACCACTCTC	1778
QY	1422	CCATGTACAAATCCAGTGTTCCTATGAACAATCGCCCAATAGTATGATCAAAACGGATGA	1481
Db	1779	TCATGTACAAATTCATCTACCCAATCCACAAAGGCTTTGATTTGTTTGGCACTG	1838
QY	1482	ATTATCAATTTACAAATTTGTCGACAGGTGATGACAGAGATGGACAGATGATGATG	1541
Db	1839	ACTACAAGTATACAAAGATAGCTGTGGATCGAGTGAACGCTGCTGATGGGAGATACCATG	1898
QY	1542	TTATGTTTATCGGAACAGATGTGGGACCGTTCTTAAAGTAGTTTCAATTTCTTAAGGAGA	1601
Db	1899	TCCTGTTTCTCGGAACAGATCGGGGTCTGTGCAAAAGTGTGTTGTTCTTCTACTACA	1958
QY	1602	CTTGATGTATTTAGAAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGAACCGCATG	1661
Db	1959	---ACTCTGTCAGTGGGAGCTCATCTCGAGGAGCTGGAAGTCTTTTAAAGATCATGCTC	2015
QY	1662	CTATTTCAGCAATGGAGCTTTCCATAGAGAGCAACATATATATTGGTTCAAGCGCTG	1721
Db	2016	CTATAACAACAATGAAATTTTCATCTAAAGAGCAACAGTTGTATGTAGTGTCCAAATGAAG	2075
QY	1722	GGTTTGGCCAGCTCCCTTTTACACCGTGTGATATTTAGGGAAGACGCTGCTGAGTGT	1781
Db	2076	GGGTTTCCAGGATATCTGCAACCGCTGCCACATCTATGGTACAGCTTGTGCTGACTGCT	2135
QY	1782	GCCTCGCCGAGACCTTTACTGTCTGGATGTTCTTGCATGTTCTCGCTATTTTCCCA	1841
Db	2136	GCCTGGCGGGACCTTTATTTGGCTGGATGGCCATCTCTGTTCCAGATTTCTACCCAA	2195
QY	1842	CTGCAAGAGACGCAAGAGCAAGATATAAGAAATGGAGACCCACTGCTCACTGTT	1901

Db 2196 CTGGAAACGGAGGCGCGAAGACACAGATGTGAGACATGGAACCCACTGACTCAATGCA 2255
Qy 1902 CAGACTTACACCATGATATACCAATGCGCCACAGCCCTGAAGAGAGAAATCATCTATGGTG 1961
Db 2256 GAGGATTT-----TAATCTAAAGCATACAGAAATGCAGCTGAAATTTGGCAGTATGGAG 2309
Qy 1962 TAGAAGATAGTAGCACAATTTTGGAAATGCGAGTCCGAAAGTCGAGAGAGCGCTGGTCTATT 2021
Db 2310 TAAAAAATAACACCACTTTCTGGAGTGTGCCCCCAAGTCTCCGACGAGCATCTATCAAGT 2369
Qy 2022 GCGAATTCAGAGGCGGAATGAAGAGCGGAAGAGAGAGATCAGAGTGGATGATCATATCA 2081
Db 2370 GCGCTGTATTACAGA---AAGACAAGACAGGAGGAAGAGGTTAAGCTGGAATGAACGAATAA 2426
Qy 2082 TCAGACACAGATCAAGGCTTCTGCTACGTAGTCTACAACAGAGAGATTTCAGGCAATTACC 2141
Db 2427 TAGCCACTTCACAGGACTTCCTGATCCGCTCTGTTTCAGGGTCTTACCAAGGACTTTATC 2486
Qy 2142 TCTGCCATCGGTGGAAACATGGGTTTCATACAAACTCTTCTTAAGGTAACCCCTGGAAGTCA 2201
Db 2487 ACTGCATTCTACAGAAATAGTTTCAAGCAGACCATACCCAGATCAACTTCAAGTTT 2546
Qy 2202 TTGACACAGAGCAATTTGG 2219
Db 2547 TAGATTCAAGAAATGGTGG 2564

RESULT 6

US-10-357-930-28973
; Sequence 28973, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28973
; LENGTH: 5231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2..5231
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28973

Query Match 20.5%; Score 556; DB 6; Length 5231;
Best Local Similarity 57.6%; Pred. No. 2.5e-168;
Matches 1117; Conservative 0; Mismatches 800; Indels 21; Gaps 6;
Qy 288 CAAGCTGAAATATTCCTACAAGAAATGTTGGAATCCCAATGATGATCACTTTCAATG 347
Db 642 CAAGAGTTTATTAAACATTTGATGAACCTTCGAGAAACCAAGACCTCTGAAATACCTTCAGCC 701

Qy 348 GCTTGGCCAAACAGCTCCAGTTATCATACCTTCTCTTTTGGATGAGGAACGAGTAGCTGT 407
Db 702 TTTCCCACTCTCTTTAGACTACAGGATTTTATTAATGATGAAGATCAGGACCGGATAT 761
Qy 408 ATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATCA---AGGATTTTC 464
Db 762 ATGTGGGAAGCAAGATCACATCTCTTCCCTGAAATATTAAACAATATAAGTCAAGAAGCTT.821
Qy 465 AAAGATTTGTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGCTGGGAAAG 524
Db 822 TGAGTGTCTTCTGGCCAGCATCTACAATCAAGTTGAAAGATGCAAAATGGCTGGCAAG 881
Qy 525 ACATCTCTGAAAGATGTCTAAATTTTCATCAAGGTACTTTAAGGCATATATATCAGATCTACT 584
Db 882 ATCCACACACAGCTGTGGGAACCTTTGTCCGTGTAATTCAGACTTTCATCGCACATTT 941
Qy 585 TGTAGCCTGTGGAACGGGGCTTTTTCATCAATTTGCACTTACCTACATTTGAAATTTGACATC 644
Db 942 TGTATGTCTGTGGAGTGGCGCTTTCAGTCTCTGTACTTACTTGAACAGAGGAGGA 1001
Qy 645 ATCTTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACACCGCCGTGGGAAGA 704
Db 1002 GATCAGAGGACCAAGTTTTCATGATTG---ACTCCAAGTGTGAATCTCGAAAGACGCGT 1058
Qy 705 GTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAAATTTATATCTGT 764
Db 1059 GCTCTTTCAACCCCAACGTAACCGGTGCTGTTATGATCAATGAGAGGCTTTTCTCTG 1118
Qy 765 GAACTGACGTGATTTTATGGGGCGAGACTTTTGTCTATCTCCGAACCTTTGGGCAACACC 824
Db 1119 GAATGTATATAGATTTTCATGGGACAGATGCTGCTATTTTTCGAAGTTTAAACCAAGAGA 1178
Qy 825 ACCCAATCAGACAGAGCAGCATGATTCAGGCTGGCTCAATGATCCAAAGTTTCATTAGTG 884
Db 1179 ATGCGGTGAGAACTGATCAACATAATTCCAATGGCTAAGTGAACCTATGTTTGTAGATG 1238
Qy 885 CCCACCTCATCTCAGAGAGTGACAATCTGAAGATGACAAAAGTATATCTTTTCTTCGCTG 944
Db 1239 CACATGTCATCCAGATGGTACTGATCCAAATGATGCTAAGTGTACTTCTTCTTCAAG 1298
Qy 945 AAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCAGCTAGATAAGTTCAGATAT 1004
Db 1299 AAAAATGACTTTGGAGGCGACAGAGTCTGCTGAATAAATGGAACACATTTCTCTCAAG 1064
Qy 1005 GCAAGATGACTTTGGAGGCGACAGAGTCTGCTGAATAAATGGAACACATTTCTCTCAAG 1064
Db 1359 GTCCCTAATGACACTGGTGGACTGCTGAGCTTGTCAACAAGTGGACCACTTTCTTAAAGG 1418
Qy 1065 CTGCTCTGATTTGCTCAGTGCAGTCCAAATGGCAATTGACACTCATTTTGTATGAATGTC 1124
Db 1419 CGAGCTGGTGTGCTCGGTAACAGATGAAGACGCGCCCAAGAAACACACTTTGATGAATAG 1478
Qy 1125 AGGATGATTCCTAATGAACCTTTTAAAGATCCTAAAAATCCAGTTGTATATGAGAGTTTAA 1184
Db 1479 AGGATGTTTCTGCTGGAACCTGATAACCCGAGGACAACTAGTGTATGGCATTTTAA 1538
Qy 1185 CGACTTCAGTAAACATTTTCAAGGGATCAGCCGTGTGTATGATAGCATGATGATGTA 1244
Db 1539 CAACATCAAGCTCAGTTTTCAAAGGATCAGCCGTGTGTGTATCATTTATCTGTATATAC 1598
Qy 1245 GAAGGCTGTTCTTGGTCCATATATCCACAGGATGGACCCCAACTATCAATGGGGTGCCTT 1304
Db 1599 AGACTGTGTTTAAATGGGCTTTTGGCCCAAGAAAGGGCCCAATCATCAGCTGATTTCT 1658
Qy 1305 ATCAAGGAAGAGTCCCTTATCCAGGCGCAGGAACCTTGTCCAGCAAAAACATTT---GGTG 1361
Db 1659 ATCAGGCGAGAAATTCATATCTCGCCCTGGAACCTTGTCCAGGAGGAGCAATTTACACCA 1718
Qy 1362 GTTTTGTACTTCAAAAGGACCTTCTGATGATGTTTAAACCTTTTGCAGAGAGTCAATCCAG 1421
Db 1719 ATATGCGAAACCAAGGAGTTCAGATGATGTTGTCACTTTTATTTATTCGGAACCATCTC 1778

Qy	1422	CCATGTACAATCCAGTGTTCCTTATGAACAATCGCCCAATAGTGAATCAAAACGGATGTAA	1481
Db	1779	TCATGTACAATTCATCTACCCATCACAAAAGGCGCTTGATGTGTTTCGTATTTGGCAGCTG	1838
Qy	1482	ATTATCAATTTACACAAATTGCTGACACGAGTGGATGCAGAAGATGGACAGTATGATG	1541
Db	1839	ACTACAGTATACAAAGTAGCTGTGATCGATGTAACGCTGCTGATGGAGATACCATG	1898
Qy	1542	TTATGTTTATCGGAACAGATGTTGGGACCGTCTTTAAAGTAGTTCCTCAATTCCTTAAGGAGA	1601
Db	1899	TCCTGTTTCTCGGAACAGATCGGGGTACTGTGCAAAAGTGGTGTGTTCTTCTCTACTAAACA	1958
Qy	1602	CTTGGTATGATTAGAGAGAGTTCGTCTGGAAGAAATGACAGTTTTTCGGGAACCGAGCTG	1661
Db	1959	---ACTCTGTGAGTGGCGAGCTCAATCTGAGGAGCTGGGAAGTCTTTAAGAAATCATGCTC	2015
Qy	1662	CTATTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATTGGTTCACACGGCTG	1721
Db	2016	CTATACAACAATGAAATTTTCATCTAANAAGCAACAGATTGTATGTGAGTTCCAATGAAG	2075
Qy	1722	GGGTTGCCAGCTCCCTTTACACCGGTGTGATATTATACGGGAAGCGTGTGCTGAGTGTT	1781
Db	2076	GGGTTTCCCAGGTATCTCTGCACCGCTGCCACATCTATGGTACAGCGTGTGCTGCACTGCT	2135
Qy	1782	GCCTCGCCCCAGACCCCTTACTGTGCTTGGGATGGTTCGTGATGTTCTCGCTATTATTCCCA	1841
Db	2136	GCCTTGGCGGGAGCCCTTATTGGCCCTGGGATGCCATTCCTGTTTCAGAAATCTACCCAA	2195
Qy	1842	CTGCAAAAGAGACGCACAAGACGACAAATATAAGAAATGGAGACCCACTGACTCACTGTT	1901
Db	2196	CTGGAAACGGAGGCGGAAGACAAGATGTGAGACATGGAACCCCACTGACTCAATGCA	2255
Qy	1902	CAGACTTACCATGATTAATCAACATGGCCACAGCCCTGAAGAGAGAAATCATCTATGGTG	1961
Db	2256	GAGGATTT-----TAATCTAAAGCATACAGAAATGCAGCTGAAATTTGTCAGATATGGAG	2309
Qy	1962	TAGAGAATAGTAGCACATTTTGGAAATGCAGTCGGAAGTCGCAAGAGCGCTGGTCTATT	2021
Db	2310	TAAANAATAACACATTTTCTGGAGTGTGCCCCCAAGTCTCCGAGGCATCTATCAAGT	2369
Qy	2022	GGCAATTCGAGGCGAAATGAAGAGCGAAAGAAAGAGATCAGAGTGGATGATCATATCA	2081
Db	2370	GGCTGTTTACAG---AAGACAAGACAGGAGGAAGAGGTTAAGCTGAATGAACGAATAA	2426
Qy	2082	TCAGGACAGATCAAGGCTTCTGCTAGTGTCTTAACAAGAGGATTCAGGCAATTAACC	2141
Db	2427	TAGCCACTTCACAGGGACTCTGATCCGCTCTGTTCAGGGTTCCTGACCAAGGACTTTATC	2486
Qy	2142	TCTGCCATGCGTGGAAACATGGGTTCAATAAACTCTTCTTTAAGGTAACCCCTGGAAAGTCA	2201
Db	2487	ACTGCATTTGTACAGAAANAAGTTTTCAAGCAGACCATAGCCAGATCACTTCAGAGTTT	2546
Qy	2202	TTGACACAGAGCAATTTGG	2219
Db	2547	TAGATTCAGAAATGGTGG	2564

RESULT 7

```

US-101-357-820-27
; Sequence 27, Application US/10357820
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Casman, Stacie J.;
; APPLICANT: Gorman, Linda;
; APPLICANT: Ji, Weizhen;
; APPLICANT: Kekuda, Ramesh;
; APPLICANT: Li, Li;
; APPLICANT: Padigaru, Muralidhara;
; APPLICANT: Patturajan, M.A.;
; APPLICANT: Pena, Carol E.A.;
; APPLICANT: Shenoy, Suresh G.;
; APPLICANT: Shinkets, Richard A.;

```

```

1  APPLICANT: Stone, David J.;
2  APPLICANT: Taupier, Raymond J.
3  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METEOD
4  FILE REFERENCE: 21402-538B
5  CURRENT APPLICATION NUMBER: US/10/357,820
6  CURRENT FILING DATE: 2003-02-03
7  PRIOR APPLICATION NUMBER: 09/679460
8  PRIOR FILING DATE: 2000-10-04
9  PRIOR APPLICATION NUMBER: 09/730617
10 PRIOR FILING DATE: 2000-12-05
11 PRIOR APPLICATION NUMBER: 10/074978
12 PRIOR FILING DATE: 2002-02-12
13 PRIOR APPLICATION NUMBER: 10/138588
14 PRIOR FILING DATE: 2002-05-01
15 PRIOR APPLICATION NUMBER: 60/387002
16 PRIOR FILING DATE: 2002-06-07
17 PRIOR APPLICATION NUMBER: 60/355099
18 PRIOR FILING DATE: 2002-02-08
19 PRIOR APPLICATION NUMBER: 60/375579
20 PRIOR FILING DATE: 2002-04-25
21 PRIOR APPLICATION NUMBER: 60/393265
22 PRIOR FILING DATE: 2002-07-02
23 PRIOR APPLICATION NUMBER: 401825
24 PRIOR FILING DATE: 2002-08-07
25 PRIOR APPLICATION NUMBER: 60/381666
26 PRIOR FILING DATE: 2002-05-17
27 Remaining Prior Application data removed - See File Wrapper or PALM.
28 NUMBER OF SEQ ID NOS: 95
29 SOFTWARE: Curasequest version 0.1
30 SEQ ID NO 27
31 LENGTH: 1671
32 TYPE: DNA
33 ORGANISM: Homo sapiens
34 FEATURE:
35 NAME/KEY: CDS
36 LOCATION: (85) .. (1626)
37 IS-10-357-820-27

```

Query Match	15.1%;	Score 407.8;	DB 6;	Length 1671;
Best Local Similarity	57.4%;	Pred. No. 8.7e-121;		
Matches 815;	Conservative 0;	Mismatches 592;	Indels 12;	Gaps 4
Qy	288	CAAGGCTGAATTTATCTACAAAGAAATGTTGGAAATCCAAACAATGTGATCATCTTTCAATG	347	
Db	164	CAAGAGTTTATTTAAACATTTTGATGAACCTTCGAGAAACCAAGACCTCTGAAATACCTTCAGCC	223	
Qy	348	GCTTGGCCACACAGCTCCACGTTATCATACCTTTCCTTTTGGATGAGGAAACGGAGTAGGCTGT	407	
Db	224	TTTCCCAACCATCTTTTAGACTACAGGAATTTTATTAAATGATGAAGATCAGAGCCGGATAT	283	
Qy	408	ATGTTGAGCAAAAGGATCACATATTTTCAITTCGACCTGGTTAAATATCA--AGGATTTTC	464	
Db	284	ATGTGGGAAGCAAGATCACATCTTTCCCTGAAATATTAACAATATAAGTCAAGAACTT	343	
Qy	465	AAAAGATTGTGGCCAGTATCTTTACACGAGAAGAGATGAATGCAAGTGGGCTGAAAG	524	
Db	344	TGAGTGTTTCTGGCCAGCATCTACAATCAAAGTTGAAGATGCAGAAATGCGTGCAGAAAG	403	
Qy	525	ACATCTGGAAGAATGTGCTAAATTTCAATCAAGGTACTTTAAGGCATATAATCAGACTCACT	584	
Db	404	ATCCACACACGCGTGTGGAACTTTGTCGCTGTAAATTCAGACTTTTCAATCGCACACATT	463	
Qy	585	TGTACGCGCTGTGAAACGGGGCTTTTCATCCAAATTTGCACCTACATATGAAATTCGACATC	644	
Db	464	TGTATGTCTGTGGAGTGGCGCTTTTCACTCTGTCTGTACTTCTTGAACAGAGGGAGGA	523	
Qy	645	ATCCTGAGGACAATATTTTTAAAGCTGGAGAACTCACATTTTTTGAAGAACGGCCGTGGGAAGA	704	
Db	524	GATCAGAGGACCAAGTTTTCATGATTGA--CTCCAAGTGTGATCTGGAAGAACGCT	580	
Qy	705	GTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTTTAATAGATGGAATAATATCTCTG	764	
Db	581	GCTCTTTTCACCCCAACGCTGAAACACGGTGTCTGTATGATCAATGAGAGCTTTTCTCTG	640	


```
QY 765 GAACTCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGAACTCTTGGGCAACC 824
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 GAATGTATATAGATTTTCATGGGACAGATGCTGCTATTTTTCGAAGTTTAAACAAGAGA 700
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 ACCAATCAGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTAGTG 884
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 701 ATCGGTGAGACTGATCAACATAATTCAAATGGCTAAGTGAACCTATGTTTGTAGATG 760
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 CCCACCTCATCTCAGAGAGTGACATCTCGAAGATGACAAAGTATATCTTTCTCCGTTG 944
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 761 CACATGTCATCCAGATGCTGATGCTCAATGATGCTAAGGTGACTTCTTCTTCAAG 820
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 AAAATGCAATAGATGAGAACACTCTGGAAGAGTACTCACGCTAGAAATAGTTCAGATAT 1004
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 AAAAAGTACTGACAAATACAGGAGCAGCAAGATTCATTCATGATGCTCGAATAT 880
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1005 GCAAGATCACTTTGAGGGGACAGAAAGTCTGGTGAATAAATGGCAACATTCCTCAAG 1064
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 GTCCTAATGACACTGGTGACTGCGTAGCCTTGTCAACAAGTGGACCACTTCTTAAAGG 940
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1065 CTGCTCTGATTTGCTCAGTGCAGGTCGAAATGGGATTCGACACTCATTTTGTAGTACTGC 1124
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 CGAGGTGCTGCTCGTAAACAGATGAAGACGGCCAGAAACACACTTTTGTAGTAAATAG 1000
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1125 AGGATGATTTCTTAATGAATTTAAAGATCCTAAAGATCCAGTGTGTATGAGTGTATA 1184
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 AGAATGTTTCTGCTGGAACATGATAACCGAGGACACACTAGTGTATGCAATTTTA 1060
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1185 CGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTGTATGATAGCATGATGTGA 1244
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1061 CAACATCAAGCTCAGTTTTTCAAGGATCAGCCGTGTGTATCATTTATCTGTATATAC 1120
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1245 GAAGGTGTTCTTGTTCATATGCCACAGGATGGACCCAACTATCAATGATGGTGCCTT 1304
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1121 AGACTGTGTTTAAATGGGCTTTTGGCCACAAAGAGGGGCCAACTCATCAGCTGATTCCT 1180
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1305 ATCAAGGAAGAGTCCCTATCCAGGCGCAGGAACCTTGTCCCAAGAAACATTTT--GGTG 1361
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1181 ATCAGGCGAGATTCATATCTCGCCCTGGAACTTGTCCAGGAGGAGATTTACACCA 1240
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1362 GTTTTGACTCTAAGAGGACCTTCTCTGATGATGTTTATACCTTTGCAAGATCATCCAG 1421
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 ATATGGAACCAAGGAGTCCACAGATGATGTTGTCACTTTTATCGGAACCATCTC 1300
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1422 CATGTACAAATCAGTGTCTTATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAA 1481
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1301 TCATGTACAAATTCATCTACCAATCCACAAAGGCTTTGATTTGTTCTATTGGCACTG 1360
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1482 ATTATCAATTTACAAAATTTGCTAGACCGAGTGGATGCAGAGATGGACAGTATGATG 1541
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1361 ACTACAGTATACAAAGATGCTGTGGATCGAGTGAACCTGCTGATGGAGATACCATG 1420
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1542 TTATGTTTATCGGAACAGATGTTGGGACCGTTCTTAAAGTACTTTTCAATTTCTTAAGGAGA 1601
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1421 TCTGTTTCTCGGAACAGATCGGGTACTGTGCAAAAAGTGTGTTCTTCTACTAACA 1480
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1602 CTTGGTATGATTTAGAGAGTTCTGCTGGAAGAAATGACAGTTTTTTCGGGAACCGACTG 1661
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1481 ---ACTCTGCTAGTGGGAGCTCATTTCTGGAGGAGCTGGAAGTCTTTTAAAGATCATGCTC 1537
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1662 CTATTTTCAGCAATGGAGCTTTTCCACTAAGCAGCAAC 1700
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1538 CTATACACAAATGAATTTTATCTTAAAGAGGAGC 1576
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

```
PCT-US03-09929-25
; Sequence 25, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
```

```
; CURRENT APPLICATION NUMBER: PCT/US03/09929
```

```
; CURRENT FILING DATE: 2003-04-01
```

```
; PRIOR APPLICATION NUMBER: 60/368,996
```

```
; PRIOR FILING DATE: 2002-04-01
```

```
; PRIOR APPLICATION NUMBER: 60/369,980
```

```
; PRIOR FILING DATE: 2002-04-04
```

```
; PRIOR APPLICATION NUMBER: 60/370,381
```

```
; PRIOR FILING DATE: 2002-04-05
```

```
; PRIOR APPLICATION NUMBER: 60/370,969
```

```
; PRIOR FILING DATE: 2002-04-08
```

```
; PRIOR APPLICATION NUMBER: 60/371,002
```

```
; PRIOR FILING DATE: 2002-04-09
```

```
; PRIOR APPLICATION NUMBER: 60/372,002
```

```
; PRIOR FILING DATE: 2002-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/384,297
```

```
; PRIOR FILING DATE: 2002-05-30
```

```
; PRIOR APPLICATION NUMBER: 60/386,816
```

```
; PRIOR FILING DATE: 2002-06-07
```

```
; PRIOR APPLICATION NUMBER: 60/389,123
```

```
; PRIOR FILING DATE: 2002-06-13
```

```
; PRIOR APPLICATION NUMBER: 60/402,207
```

```
; PRIOR FILING DATE: 2002-08-09
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 179
```

```
; SOFTWARE: CuraSeqIst version 0.1
```

```
; SEQ ID NO 25
```

```
; LENGTH: 1438
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
; FEATURE:
```

```
; NAME/KEY: CDS
```

```
; LOCATION: (2)..(1438)
```

```
; PCT-US03-09929-25
```

```
Query Match 4.7%; Score 126.6; DB 1; Length 1438;
```

```
Best Local Similarity 50.6%; Pred. No. 4e-30;
```

```
Matches 489; Conservative 0; Mismatches 439; Indels 39; Gaps 6;
```

```
QY 689 AACGCCGTGGAGAGTCCATATCACCTTAAGCTGCTGACAGCATCCCTTTTATAGAT 748
```

```
DB 434 AGCGAATGCGCAGATGCCCATATGATGCCAAACATGCAACGTTGCGACTGTTTCGAGAT 493
```

```
QY 749 GGAGAAATATATCTCTGGAACCTGCGAGCTGATTTTATGGGCGGAGACTTTTGTCTATCTTCCGA 808
```

```
DB 494 GGAATACTATCTAGCCACAGTACTGACTTCTTCCATTTGACGAGTCAATTTACCGG 553
```

```
QY 809 ACTCTTGGGCAACCAACCCCAATCAGGACAGAGCATGATTCAGGTGGCTCAATGAT 868
```

```
DB 554 AGTCTTGGAGAAAGCCCTTACCTCGGACCGCTCAAGCAGCATTCAAAATGGTTGAAAGAA 613
```

```
QY 869 CCAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACATCTCGAAGATGCACAAAGTA 928
```

```
DB 614 CCATCTTTGTTCAAGCGGTGGA-----TTACGGAGATTATATC 652
```

```
QY 929 TACTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAAAGCTACTCACGCT 988
```

```
DB 653 TACTTCTTCTCAGGAAATAGCAGTGGAGTATTAACCATGSGAAGTAGTTTTCCTCA 712
```

```
QY 989 AGAATAGTCAAGATATGCAAGATGACTTTGGAGG---GCACAGAGTCTGGTGAATAA 1045
```

```
DB 713 AGAGTGGCTCAGGTTTGTAGAATGATATGAGGAGATCTCAAGAGTCTCTGGAGAAACAG 772
```

```
QY 1046 TGGACACATCTCCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCACAAATGGCATTCAC 1105
```

```
DB 773 TGGACGCTGTTCTTGAAGGCGCGTTGAATTTGCTCAGTTCTCTGG-----AGACTCTCAT 826
```

```
QY 1106 ACTCATTTTGTAGTGAAGTGCAGGATGATTTCCCTAATGAACCTTTTAAAGATCTCTTAAATCCA 1165
```

```
DB 827 TTTTATTTCAACATTTCCAGGCGAGTTACAGATGTGATTCGTATCAACGGGCGTAT--- 883
```

```
QY 1166 GTTGTATAGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGAGATACGCGGTGTATG 1225
```

```
DB 884 GTTCTCTGGCAAGCTTTTCTACACCTTTATACAGCATCCCTGGGTCTGCGAGTCTGTGCC 943
```


Query Match	4.6%;	Score 125;	DB 1;	Length 1327;
Best Local Similarity	50.5%;	Pred. No. 1.3e-29;		
Matches 488;	Conservative 0;	Mismatches 440;	Indels 39;	Gaps 6;

QY	689	AACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATAGAT	748
Db	323	AGCGGAATGCCAGATGCCCATATGATGCCAAACATGCCAACGTTGTCAGCTGTTTCGAGAT	382
QY	749	GGAGAATTACTCTCGAACTGCGAGCTGATTTTATGGGGCGAGACTTTTGCTATCTTCCGA	808
Db	383	GGAAACTATCTCAGCCACAGTGACTGACTTCTCTGCCATTGACGCAGTCATTTACCGG	442
QY	809	ACTCTTGGGCAACACCAACCCAACTCAGACACAGCAGCATGATTCACAGTGGCTCAATGAT	868
Db	443	AGCTTTGGAGAAAGCCCTACCTCGGACCGTCAAGCAGCATTCAAATGTTGGAAGAA	502
QY	869	CCAAAGTTCAATTAGTGGCCCACTCATCTCAGAGAGTGACAATCTGGAAGATGACAAAGTA	928
Db	503	CCATACTTTGTTCAAAGCCGTGGA-----TTACGGAGATTATATC	541
QY	929	TACTTTTCTTCGTGAAATGCNATAGATGGAGACACTCTGAAAAGCTACTCAGCT	988
Db	542	TACTTCTCTTCACGGGAAATAGCAGTGGAGTATAACACCATTGGAAAGGTAGTTTTCCCA	601
QY	989	AGAATAGGTCAGATATCAAGAATGACTTTGGAGG---GCACAGAAGCTCTGGTGAATAA	1045
Db	602	AGAGTGCTCAGGTTGTGAAGATGATATGGAGGATCTCAAGAGTCTCTGGAGNAACAG	661
QY	1046	TGGACAACTTCCTCAAAGTCGTCTGATTTGCTCAGTGGCAGGTCCAAATGGCAATTGAC	1105

[illegible]

Db	662	TGGACGCGTTCCTGAAGGCGCGTTGAAC	CTGCTCAGTTCCTGG-----AGACTCTCAT	715
Qy	1106	ACTCAITTTTCATGAACACTGCAGGATGTATTCCTAATGAACCTTTAAAGATCCTAAAAATCCA		1165
Db	716	TTTTATTTCACATTTCTCCAGGCAGTTACAGATGTGATTCGTATCAACGGGCGTGAT---		772
Qy	1166	GTTGTATATCGGAGTGTTTTACGACTTC	CAGTAAACATTTTCAAGGATCAGCCGCTGTGTATG	1225
Db	773	GTTGTCTCGCAACGTTTCTTACACCTTTAAACAGCATCCTCGGCTCGCAGTCTGTGGCC		832
Qy	1226	TATAGCATGAGTCGATGTGAGAAGGGTTCCTTGGTCCATATGCCACACGGATGGACCC		1285
Db	833	TATGACATGCTTGACATTCGCCAGTGT	TTTTTACTGGGAGATTCAAGNAACAGAACTCTCCT	892
Qy	1286	AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCCTATCCAGCGCCAGGAACCTTGT		1342
Db	893	GATTCACCTGGACACCCAGTTCCTGATGAACGAGTTCTTAAGCCCGACGGTCGTGT		952
Qy	1343	CCAGCMAA---ACATTTGGTGGT	TTTTTGACTCTACAAAGGACCTTCTCGTATGATGTTATA	1399
Db	953	GCTGGCTCATCTCTCTTAGAAGATATGCAACCTCCAATGAGTTCCTCGTATGATACCCCTG		1012
Qy	1400	ACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAATCGCCCA		1459
Db	1013	AACTTCATCAAGACGACCCGCTCATGGATGAGGCAGTGCCTTCATCTTCAACAGGCCA		1072
Qy	1460	ATAGTGATCAAAACGGATGTAAATATCAATTTACAAATTCGTCTGTAGACCGAGTGGAT		1519
Db	1073	TGGTTCTGTGAGAACATGGTCAGATACCGGCTTTACCAAAATTCAGTGGACACAGCTGTCT		1132
Qy	1520	GCAGAAGATGGACAGTATGATGTATGTATTATCGGAACAGATGTTGGGACCGTTCCTAAA		1579
Db	1133	GGGCCATATCAGAATCACACTGTGGTTTTTCTGGGATCAGAGAAGGAAATCATCTTGAAG		1192
Qy	1580	GTAGTTTCAATTCCTAAGAGACTCTGGTATGATTTAGAAGAGGTTCTGCTGGGAAGAAATG		1639
Db	1193	TTTTTGGCCAGATAGGAANATGTGGTTTTCTTAATGACAGCCCTTTTCTGGAGGAGATG		1252
Qy	1640	ACAGTTT	1646	
Db	1253	AGTGT	1259	

NUMBER OF SEQ ID NOS: 179
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 23
LENGTH: 1492
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1492)
PCT-US03-09929-23

Query Match 4.6%; Score 125; DB 1; Length 1492;

Best Local Similarity 50.5%; Pred. No. 1.3e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

QY 689 AACGGCCGTTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGAT 748
DB |||||
QY 488 AGCGGAATGCCAGATGCCATATGATGCCAATGATGCCAAGTGGCTGCTGAT 547
DB |||||
QY 749 GGAGAAATATATCTGGAACATGCGAGCTGATTTTATGGGGCGAGACTTTGCTATCTCCGA 808
DB |||||
QY 548 GGAAATATATCTGACGCCACAGTACTGCTTCTTCCATTTGACGCGAGTCAATTTACCGG 607
DB |||||
QY 809 ACTCTTGGSCACCAACCCATCAGGACAGCAGCATGATTCAGGTGGCTCAATGAT 868
DB |||||
QY 608 AGTCTTGGGAAGCCCTACCTCGGACCGTCAAGCAGCATGCAAAATGGTTGAAGAA 667
DB |||||
QY 869 CCAAAATTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCTTGAAGATGACAAGTA 928
DB |||||
QY 668 CCATCTTTGTTCAAGCCGTGA -----TTACGGAGATTATATC 706
DB |||||
QY 929 TACTTTTCTTCGGTGAATGCAATAGATGAGAGCACTCTGGAAGACTACTCACGCT 988
DB |||||
QY 707 TACTTCTTCTTCAGGGAATAGCAGTGGAGTAAACACCATGGGAAGGTAGTTTCCCA 766
DB |||||
QY 989 AGAATAGTTCAGATGCAAGATGACITTTGAGG---GCACAGAGTCTGATGATGATA 1045
DB |||||
QY 767 AGAGTGGCTCAGGTTTGAAGAATGATATGGAGGATCTCAAGAGTCTCTGGAGAAACAG 826
DB |||||
QY 1046 TGGCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGCATGAC 1105
DB |||||
QY 827 TGGAGTCTGCTCTGAGCGCGCTGACTGCTCAGTCTCTG-----AGACTCTAT 880
DB |||||
QY 1106 ACTCAATTTGATGAATGAGGATGATTTCTTAATGAATTTAAAGATCTTAAATCCA 1165
DB |||||
QY 881 TTTTATTTCAACATTTCTCAGGAGTACAGATGATGATTCATCAACGGCGTGAT--- 937
DB |||||
QY 1166 GTTGATATGAGTGTGTTACGACTTCCAGTACATTTTCAAGGATCAGCGGTGATG 1225
DB |||||
QY 938 GTTGCTGCGCAACGTTTCTACCTTATACAGCATCCCTGGGTCTGCACTCTGTGCC 997
DB |||||
QY 1226 TATAGCATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCACAGGATGGACC 1285
DB |||||
QY 998 TATGATGCTTGATGATTCAGCTTCCAGTACATTTTCAAGGATCAGCGGTGATG 1057
DB |||||
QY 1286 AACTATCAATGGTGCC---TTATCAAGGAAGAGTCCCTATCCAGGCGCAGAACTTGT 1342
DB |||||
QY 1058 GATTCACCTGGACACCACTTCTGATGAACGAGTCTTAAAGCCAGGCGAGGTTGCTGT 1117
DB |||||
QY 1343 CCCAGCAA---ACATTTGGTGGTTTGAATCTTCAAGGACCTTCTGATGATGATATA 1399
DB |||||
QY 1118 GCTGGCTATCTCTCTTGAAGAAGATGCAACCTCAATGATGTTCCCTGATGATACCTGT 1177
DB |||||
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCATGATCAATTCAGTGTTCCTTATGAACATCGCCA 1459
DB |||||
QY 1178 AACTTCAATCAAGACGACCGCTCATGATGAGGAGTGGCTCCCTCATCTTCAACAGGCA 1237
DB |||||
QY 1460 ATAGTCAATCAACCGATGATAAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
DB |||||
QY 1238 TGGTCTCTGAGAACAAATGTCAGATACCGCTTACCAAATTTGAGTGGACACACTGCT 1297
DB |||||
QY 1520 GCAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAAA 1579
DB |||||

DB 1298 GGGCCATATCAGATCACACTGTGTGTTTTCTGGATCAGAGAGGGAATCATCTTGAAG 1357
QY 1580 GTAGTTTCAATTTCTTAAAGAGACTTGGTATGATTTAGAGAGGTTCTCTGGAAGAAATG 1639
DB |||||
QY 1358 TTTTGGCCAGATAGGAATAGTGTGTTTCTAAATGACAGCCTTTTCTGGAGAGATG 1417
DB |||||
QY 1640 ACAGTTT 1646
DB |||||
QY 1418 AGTGT 1424

RESULT 12

PCT-US03-09929-33

Sequence 33, Application PC/TUS0309929

GENERAL INFORMATION:

APPLICANT: Curagen Corporation, et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-573B-061

CURRENT APPLICATION NUMBER: PCT/US03/09929

CURRENT FILING DATE: 2003-04-01

PRIOR APPLICATION NUMBER: 60/368,996

PRIOR FILING DATE: 2002-04-01

PRIOR APPLICATION NUMBER: 60/369,980

PRIOR FILING DATE: 2002-04-04

PRIOR APPLICATION NUMBER: 60/370,381

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: 60/370,969

PRIOR FILING DATE: 2002-04-08

PRIOR APPLICATION NUMBER: 60/371,002

PRIOR FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 60/372,002

PRIOR FILING DATE: 2002-04-12

PRIOR APPLICATION NUMBER: 60/384,297

PRIOR FILING DATE: 2002-05-30

PRIOR APPLICATION NUMBER: 60/386,816

PRIOR FILING DATE: 2002-06-07

PRIOR APPLICATION NUMBER: 60/389,123

PRIOR FILING DATE: 2002-06-13

PRIOR APPLICATION NUMBER: 60/402,207

PRIOR FILING DATE: 2002-08-09

Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 179

SOFTWARE: CuraseqList version 0.1

SEQ ID NO 33

LENGTH: 1878

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1878)

PCT-US03-09929-33

Query Match 4.6%; Score 125; DB 1; Length 1878;

Best Local Similarity 50.5%; Pred. No. 1.5e-29;

Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

QY 689 AACGGCCGTTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGAT 748
DB |||||
QY 424 AGCGGAATGCCAGATGCCATATGATGCCAATGATGCCAAGTGGCTGCTGAT 483
DB |||||
QY 749 GGAGAAATATATCTGGAACATGCGAGCTGATTTTATGGGGCGAGACTTTGCTATCTCCGA 808
DB |||||
QY 484 GGAAATATATCTCAGCCACAGTACTGCTTCTTGGCAATGACGAGTCAATTTACCGG 543
DB |||||
QY 809 ACTCTTGGSCACCAACCCATCAGGACAGCAGCATGATTCAGGTGGCTCAATGAT 868
DB |||||
QY 544 AGTCTTGGGAAGCCCTTACCTCGGACCGTCAAGCAGCATGCAAAATGGTTGAAGAA 603
DB |||||
QY 869 CCAAAATTCATTTAGTGGCCACCTCATCTCAGAGAGTGACATCTTGAAGATGACAAGTA 928
DB |||||
QY 604 CCATCTTTGTTCAAGCCGTGA -----TTACGGAGATTATATC 642
DB |||||
QY 929 TACTTTTCTTCCGTGAAATGCAATAGATGAGAAACACTCTGGAAAGACTACTCACGCT 988
DB |||||

Db 643 TACTTCTCTTCAGGAAATAGCAGTGGAGTATACACCATCGGAAAGGTAGTTTCCCA 702
Qy 989 AGAATAGTCAGATATGCAAGATGACTTTGGAGG---GCACAGAGTCTGTGTGAATAA 1045
Db 703 AGAGTGGCTCAGGTTTGAAGATGATATGGGAGGATCTCAAGAGTCTCGGAAACAG 762
Qy 1046 TGGACAACTTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCAAATGSCATTGAC 1105
Db 763 TGACGCTGTTCTGAAGGCGCTTGAACCTGCTCAGTTCCTGG-----AGACTCTCAT 816
Qy 1106 ACTCATTTTGTATGAACCTGAGGATGATTTCTTAATGAATTTAAAGATCCTAAATCCA 1165
Db 817 TTTTATTTCAACATTCTCCAGGAGTTACAGATGATGATTCGTATCAACGGGCGTGAT--- 873
Qy 1166 GTTGATATGAGAGTGTATGAGTCTCCAGTACATTTTCAAGGATCAGCGGTGTGTATG 1225
Db 874 GTTGTCTCGCAACGTTTCTACACCTTATAACAGCATCCCTGGGTCTGCACTCTGTGCC 933
Qy 1226 TATAGCATGATGATGAGAGGTTTCTTGTGTCATATGCCACAGGATGGACCC 1285
Db 934 TATGACATGCTTGACATGCGGAGTTTCTTACTGGAGATTCAGGAACAGAGTCTCT 993
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAATTGT 1342
Db 994 GATTCCACCTGGACACCACTTCTGTATGAACGAGTTCCTAAGCCAGGCCAGTTGCTGT 1053
Qy 1343 CCAGCAAAA---ACATTTGGTGTGTTGACTCTCAAGAGGACCTTCTGATGATGTATA 1399
Db 1054 GCTGGCTCATCTCTCTTAGAAGATATGCAACCTCCAATGAGTTCCTGATGATACCCCTG 1113
Qy 1400 ACTTTGCAAGAGTATCCAGCATCCAGCATGATCAATCCAGTCTTCTTATGAACATCGCCA 1459
Db 1114 AACTTCATCAAGACGACCCGCTCATGATGAGGAGTCCCTTCAATCTTCAACAGGCCA 1173
Qy 1460 ATAGTATCAAAACGAGTGTAAATTTATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 1174 TGTCTCTGAGAACATGTCAGATACCGCTTACCAAAATTCAGTGGACACAGTGT 1233
Qy 1520 GCAGAGATGGACAGTATGATGTTATGTTTATCGGAAACAGATGTTGGGACCGTCTTAA 1579
Db 1234 GGGCCATATCAGAAATCACACTGTGGTTTCTGGATCAGAGAGGGAATCATCTTGAAG 1293
Qy 1580 GTAGTTTCAATCTCAAGCAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 1294 TTTTGGCCAGATAGGAATAGTGGTTTCTAAATGACAGCCTTTTCTGGAGGAGATG 1353
Qy 1640 ACAGTTT 1646
Db 1354 AGTGTTT 1360

RESULT 13
PCT-US03-09929-15
; Sequence 15, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curogen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 15
; LENGTH: 1921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
PCT-US03-09929-15

Query Match 4.6%; Score 125; DB 1; Length 1921;
Best Local Similarity 50.5%; Pred.No. 1.5e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

Qy 689 AACGCCGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATAGAT 748
Db 406 AGCGAATGGCCAGATGCCATATGATGCCAAACATGCCAAGCTTGCATGTTTGCAGAT 465
Qy 749 GGAGAAATPATCTCTGGAACTGACAGTGAATTTTATGGGGCGAGACTTTGTATCTTCCGA 808
Db 466 GGAAAACTATCTCAGCCACAGTACTGACTTCTTGGCCATTCAGCGAGTCAATTTACCGG 525
Qy 809 ACTCTTGGGACACACCACCAATCAGGACAGAGCATGATTCAGTGGTGGTCAATGAT 868
Db 526 AGTCTTGAGAAAGCCCTTACCCTCGGACCGCTCAAGCACGATTTCAAAATGGTTGAAGAA 585
Qy 869 CCNAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGAACAATCTCTGAAGATGACAAAGTA 928
Db 586 CCATCTTTTTCAGCGCTGGA-----TTACGGAGATTATATC 624
Qy 929 TACTTTTTCTTCGTTGAAAAATGCAATAGATGGAGAACACTCTCTGAAAAAGCTACTCACGCT 988
Db 625 TACTTCTTCTTCAGGAAATAGCAGTGGAGTATAACACCATCGGAAAGGTAGTTTTCCTCA 684
Qy 989 AGATAGTCAGATATGCAAGAAATCACTTTTCGAGG---GCACAGAGTCTGTGTGAATAA 1045
Db 685 AGAGTGGCTCAGGTTTGAAGAAATGATATGGAGGATCTCAAGAGTCTCTGGAGAACAG 744
Qy 1046 TGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAAATGGCATTTGAC 1105
Db 745 TGGAGTCTGTTCTGAGGCGGCTTGAACCTGCTCAGTTCCTGG-----AGACTCTCAT 798
Qy 1106 ACTCATTTTGTATGAATGTCAGGATGATTTCTTAATGAATTTAAAGATCTCTAAATAATCA 1165
Db 799 TTTTATTTCAACATTTCTCAGGAGTTACAGATGATTCGTATCAACGGGCGTGAT--- 855
Qy 1166 GTTGATATGAGTGTTTTACACCTTCCAGTAAACATTTTCAAGGATCAGCGGTGTGTATG 1225
Db 856 GTTGTCTCGCAACGTTTCTTACACCTTTATAACAGCATCCCTGGGTCTGCACTCTGTGCC 915
Qy 1226 TATAGCATGATGATGAGAGGTTTCTTGTGTCATATGCCACAGGATGGACCC 1285
Db 916 TATGACATGCTTGACATTTGCCAGTGTTTTACTGGAGATTCAAGGAACAGAGTCTCT 975
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACCTGT 1342
Db 976 GATTCCACCTGGACACCACTTCTGATGAACGAGTTCCTAAGCCAGGCCAGGTTGCTGT 1035
Qy 1343 CCCAGCAAAA---ACATTTGGTGGTTTGTGACTCTCAAGAGGACCTTCTCTGATGATGTATA 1399
Db 1036 GCTGGCTCATCTCTCTTAGAAGATATGCAACCTTCAATGAGTTCCTCTGATGATACCCCTG 1095
Qy 1400 ACCTTTGCAAGAGTTCATCCAGCCATGTAACAATCCAGTGTTCCTTATGAACAAATCGCCA 1459

Db 1096 AACTTCATCAAGACGCCCGCTCATGGATGAGGAGTGCCTCATCTTCAACAGGCCA 1155
Qy 1460 ATAGTGCATCAAAACCGGATGTAATATCAATTTTACACAAATTTGCTGTAGACCGAGTGGAT 1519
Db 1156 TGGTTCTTCAGAACCAATGGTCAAGTACCGCTTACCAAAATTTGCAAGTGGACACAGCTGCT 1215
Qy 1520 GCAGAGATGGACATGATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTTTAAA 1579
Db 1216 GGGCCATATCAGAAATCAGACTGTGTTTCTGGGATCAGAGAAGGAATCATCTTGAAG 1275
Qy 1580 GTAGTTTCAATTCCTAAGAGAGTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 1276 TTTTGGCCAGAAATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTCGAGGAGATG 1335
Qy 1640 ACAGTTT 1646
Db 1336 AGTGTTT 1342

RESULT 14
PCT-US03-09929-49
; Sequence 49, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 49
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1948)
PCT-US03-09929-49

Query Match 4.6%; Score 125; DB 1; Length 1948;
Best Local Similarity 50.5%; Pred. No. 1.5e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;
Qy 689 AACGCCGTGGGAAGTCCATATGACCCCTAGCTGTGTCAGACATCCCTTTTAAATAGAT 748
Db 434 AGCGGAATGGCCAGATGCCCATATGTCACCAACATGCAAGTTCGCACTGTTTCAGAT 493
Qy 749 GCAGAAATATATCTCTGGAATCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCCGA 808
Db 494 GGAAACTATATCTACGCCACAGTACTGCTTCTTGCATTTGACGAGTCAATTTACCGG 553

Qy 809 ACTCTTGGGACACACCACCCCAATCAGACAGAGAGATGATTCAGGTGCGTCAATGAT 868
Db 554 AGTCTTGAGAAAGCCCTTACCCTCGGACCCCTCAAGCAGCATCAAAATGGTTGAAAGAA 613
Qy 869 CCAAGTTTCATTTAGTGCCACCTCATCTCAGAGAGTGACATCCTGGAAGATGACAAAGTA 928
Db 614 CCATCTTTGTTCAAGCCGTGGA-----TTACGGAGATTATATC 652
Qy 929 TACTTTTTTCTTCGTTGAAAATGCAATAGATGGAGAACACTCTCGAAAAAGCTACTCACGCT 988
Db 653 TACTTCTTTCAGGGAATAGCAGTGAGATTAACACCATGSGAAGTAGTTTTCCTCA 712
Qy 989 AGAATAGTCTAGATATGCAAGAAATGACTTTTGGAGS---GCACAGAGTCTCGTGAATAAA 1045
Db 713 AGAGTGGCTCAGGTTTGTGAAGAAATGATATGGGAGGATCTCAAGAGATCTCTGGAGAACAG 772
Qy 1046 TGGACAACTTCCTCAAGCTCGTCTGATTTTGGTCTCAGTCCAGGTCCTCAAAATGGCAATGAC 1105
Db 773 TGGAGCTGTTCTTGAAGGCGGCTTGAATGCTCAGTTCTCTGG-----AGACTCTCAT 826
Qy 1106 ACTCATTTTGTGAACTGCAGGATGATTTCTTAATGAACTTTTAAAGATCTCTAAAAATCCA 1165
Db 827 TTTTATTTCAACATTTCTCCAGGAGTTACAGATGTGATTCGTATCAACGGCGGTGAT--- 883
Qy 1166 GTTGATATAGAGTGTTTTACGACTTTCAGATTAACATTTTCAAGGGATCAGCGTGTGTATG 1225
Db 884 GTTGTCTCGGCAAGTGTCTTACACCTTATAACAGCATCCCTGGGTCTGCACTCTGTGCC 943
Qy 1226 TATAGCATGATGATGAGNAGGTTGTTCTTGGTCCATATGCCACAGGATGGACCC 1285
Db 944 TATGACATGCTTGACATTTGCCAGTGTTTTACTGGGAGATTCAGGAAACAGAAAGTCTCCT 1003
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCCTATCCACGGCCAGGAACCTTGT 1342
Db 1004 GATTCACCTGGACACCAAGTTCCTGTATGAACAGAGTTCCTAAGCCAGGCGAGTGTCTGT 1063
Qy 1343 CCCAGCAAAA---ACATTTGGTGGTTTGTGACTCTCAAAAGGACCTTCTCTGATGATGTATA 1399
Db 1064 GCTGCTCATCTCTCTTAGAAAGATATGCAACCTCCAATGAGTTCCTCTGATGATACCCGTG 1123
Qy 1400 ACCTTTGCAGAAAGTCCATCCAGCCATGTACATCCAGTGTTCCTATGAACAAATGCCCA 1459
Db 1124 AACTTTCATCAAGCGCACCCGCTCATGGATGAGGAGTGCCCTCATCTTCAACAGGCCA 1183
Qy 1460 ATAGTATCAAAACCGGATGTAATATCAATTTTACACAAATTTGCTGTAGACCGAGTGGAT 1519
Db 1184 TGGTTCTCGAACAATGGTCAAGTACCGCTTACCAAAATTTGAGTGGACACAGCTGCT 1243
Qy 1520 GCAGAAATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 1244 GGGCCATATCAGAAATCAGACTGTGTTTCTGGATCAGAGAAGGAATCATCTTGAAG 1303
Qy 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
Db 1304 TTTTGGCCAGATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTCGAGGAGATG 1363
Qy 1640 ACAGTTT 1646
Db 1364 AGTGTTT 1370

RESULT 15
PCT-US03-09929-51
; Sequence 51, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980

PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: 60/402,207
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 179
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 51
LENGTH: 2583
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2583)
PCT-US03-09929-51

Query Match 4.6%; Score 125; DB 1; Length 2583;
Best Local Similarity 50.5%; Pred. No. 1.8e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

QY 689 AACGCCGTGGGAGAGTCCATATGACCCCTAGCTGCTGACAGCATCCCTTTTATAGAT 748
DB 1129 AGCGGAATGCCAGATGCCCATATGATGCCAACAATGCCAACGTTGCACTGTTGCAGAT 1188
QY 749 GGAGAAATTAATCTCTGGAATCTGAGCTGATTTTATGGGCGGAGACTTTGCTATCTTCCGA 808
DB 1189 GGAAACTATATCTAGCCACAGTACTGCTTCTTGGCCATGACGAGTCAATTTACCGG 1248
QY 809 ACTCTGGGACACACCCCAATCAGGACAGAGCAGATGATTCAGGTGGCTCAATGAT 868
DB 1249 AGCTTGGAGAAAGCCCTACCCCTGGGACCGTCAAGACAGATCAAAATGGTTGAAGAA 1308
QY 869 CCAAGTCTATTAGTGGCCACCTCATCTCAGAGATGACATCCTGAAGATGACAAAGTA 928
DB 1309 CCATACTTTGTTCAAGCCCGTGA-----TTACGGAGATTATATC 1347
QY 929 TACTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCACGCT 988
DB 1348 TACTTCTTCTCAGGGAATAGCAGTGGAGTATACCATGGGAAAGGTAGTTTCCCA 1407
QY 989 AGAATAGTTCAGATGCAAGAAATGACTTTGGAGG---GCACAGAGTCTGGTGAATAAA 1045
DB 1408 AGAGTGGCTCAGGTTTGAAGAAATGATATGGGAGGATCTCAAAGAGTCTCTGGAGAAACAG 1467
QY 1046 TGGACAACTTCTCAAGCTGCTGATTTGCTCAGTGGCAGGTCCAAATGGCATTTGAC 1105
DB 1468 TGGACGTGCTTCTGAAGCGCGCTTGAACCTGCTCAGTCTCTG-----AGACTCTCAT 1521
QY 1106 ACTCATTTTGAATCTGAGGATGATTTCTTAATGAATCTTTAAAGATCTCAAATAATCCA 1165
DB 1522 TTTTATTTCAACATCTCCAGGAGTTACAGATGATGATCTCATCAACGGGCGTAT--- 1578
QY 1166 GTTGATATGGAGTGTTCAGACTTCCAGTAAATTTTCAAGGGATCAGCCGTGTGTATG 1225
DB 1579 GTTGTCTGGCAACGTTTCTACACCTTATAACAGCATCCCTGGGTCTGCAGTCTGTGCC 1638
QY 1226 TATACCATGATGTGAGAGGGTGTCTTCTGCTCATATGCCCCACAGGGATGGACCC 1285
DB 1639 TATGATGCTTGTGACATTCGCCAGTGTCTTCTGAGGATTCAGGAAACAGAAAGTCTCT 1698

QY 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCTATCCACGCCAGGAACTTGT 1342
DB 1699 GATTCACCTGACACACAGTTCTCTGATGAACGAGTTCTTAAGCCCGAGCCAGGTTGCTGT 1758
QY 1343 CCCAGCAAA---ACATTTGGTGGTTTGTACTCTACAAAGGACCTTCTCTGATGATGTTATA 1399
DB 1759 GCTGGCTCATCTCTCTTAGAAAGATATGCAACCTCCATGATGATGATGATGATGATGATG 1818
QY 1400 ACCTTTGAAGAAGTCAATCCAGCCATGACCAATCCAGTGTCTTCTATGAACAATCGCCA 1459
DB 1819 AACTTTCATCAAGACGCCCGCTCATGATGAGGAGTGGCCCTCATCTTCAAGGCCCA 1878
QY 1460 ATAGTGATCAAAACGGATGTAAATATCAATTTTACACAAATTTGCTAGACCGAGTGGAT 1519
DB 1879 TGGTTCTTGAGAACAAATGCTCAGATACCGCTTACCAAAATTTGCAGTGGACACAGCTGT 1938
QY 1520 GCAGAGATGGACAGTATGATGTTATGTTATCGAACAGATGTTGGACCGTCTCTTAAA 1579
DB 1939 GGGCCATATCAGAATCACACTGTGTTTCTGGGATCAGAGAGGGAATCATCTTTGAAG 1998
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCTGAAGAAATG 1639
DB 1999 TTTTGGCCAGATAGGAATAGTGTCTTCTAAATGACAGCCTTTCTCTGGAGAGATG 2058
QY 1640 ACAGTTT 1646
DB 2059 AGTGT 2065

Search completed: July 31, 2003, 19:26:55
Job time : 335 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM.nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:51:05 ; Search time 181 Seconds
(without alignments)
6606.115 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatcctttatttcgatg.....aggcttttttcttaataacc 2709

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2508.4	92.6	2601	1	US-08-121-713D-53
2	2508.4	92.6	2601	1	US-08-835-268-53
3	2508.4	92.6	2601	2	US-09-060-692-53
4	2508.4	92.6	2601	3	US-08-833-391-53
5	2508.4	92.6	2601	4	US-09-060-610-53
6	2508.4	92.6	2601	5	PCT-US94-10151A-53
7	1415.2	52.2	1481	1	US-08-136-922-1
8	596.4	22.0	2898	4	US-09-308-179B-2
9	95.2	3.5	3692	4	US-09-077-940A-1
10	84	3.1	2433	4	US-09-300-958A-24
11	83.2	3.1	4157	4	US-08-556-422A-1
12	69.8	2.6	2854	1	US-08-121-713D-57
13	69.8	2.6	2854	1	US-08-835-268-57
14	69.8	2.6	2854	2	US-09-060-692-57
15	69.8	2.6	2854	3	US-08-833-391-57
16	69.8	2.6	2854	4	US-09-060-610-57
17	69.8	2.6	2854	5	PCT-US94-10151A-57
18	68.4	2.5	2790	4	US-09-254-594-5
19	68.4	2.5	3432	4	US-09-254-594-4
20	60.6	2.2	3524	4	US-09-077-940A-3
21	58.6	2.2	7218	1	US-08-232-463-14
22	58.4	2.2	3560	1	US-08-121-713D-59
23	58.4	2.2	3560	1	US-08-835-268-59
24	58.4	2.2	3560	2	US-09-060-692-59
25	58.4	2.2	3560	3	US-08-833-391-59
26	58.4	2.2	3560	4	US-09-060-610-59
27	58.4	2.2	3560	5	PCT-US94-10151A-59

28	54.8	2.0	2504	1	US-08-121-713D-63	Sequence 63, Appl
29	54.8	2.0	2504	1	US-08-835-268-63	Sequence 63, Appl
30	54.8	2.0	2504	2	US-09-060-692-63	Sequence 63, Appl
31	54.8	2.0	2504	3	US-08-833-391-63	Sequence 63, Appl
32	54.8	2.0	2504	4	US-09-060-610-63	Sequence 63, Appl
33	54.8	2.0	2504	5	PCT-US94-10151A-63	Sequence 63, Appl
34	54.6	2.0	2787	4	US-09-254-594-2	Sequence 2, Appl
35	54.6	2.0	3195	4	US-09-254-594-1	Sequence 1, Appl
36	48.6	1.8	2670	1	US-08-121-713D-61	Sequence 61, Appl
37	48.6	1.8	2670	1	US-08-835-268-61	Sequence 61, Appl
38	48.6	1.8	2670	2	US-09-060-692-61	Sequence 61, Appl
39	48.6	1.8	2670	3	US-08-833-391-61	Sequence 61, Appl
40	48.6	1.8	2670	4	US-09-060-610-61	Sequence 61, Appl
41	48.6	1.8	2670	5	PCT-US94-10151A-61	Sequence 61, Appl
42	45.4	1.7	2010	3	US-09-240-410-1	Sequence 1, Appl
43	45.4	1.7	2498	3	US-09-041-236-1	Sequence 1, Appl
44	45.4	1.7	2498	4	US-09-771-467C-1	Sequence 1, Appl
45	38.2	1.4	1818	3	US-09-041-236-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-121-713D-53
; Sequence 53, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Oeman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
US-08-121-713D-53
Query Match 92.6%; Score 2508.4; DB 1; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	193	CTGCAGCATGGGCTGCTTAACATAGGATGCTCTGTCTTTCTGGGAGTATTACTTACAGC	252
Db	9	CTGCAGCATGGGCTGCTTAACATAGGATGCTCTGTCTTTCTGGGAGTATTACTTACAGC	68
QY	253	AAGAGCAAACTATCAAAATGGGAAGAAACAATGGCCCAAGCTGAAATATCTTACAAAGA	312
Db	69	AAGAGCAAACTATCAAAATGGGAAGAAACAATGGCCCAAGCTGAAATATCTTACAAAGA	128
QY	313	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA	372
Db	129	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA	188
QY	373	TACCTTCCCTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGACAAAGGATCACATATT	432
Db	189	TACCTTCCCTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGACAAAGGATCACATATT	248
QY	433	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATCAAGTGGGCTCGAAAAGACATCTCTGAAGAAATGTCTAAATTCAT	552
Db	309	CAGAAGAGATGAATCAAGTGGGCTCGAAAAGACATCTCTGAAGAAATGTCTAAATTCAT	368
QY	553	CAAGGTACTTAAAGCATATAATCAGACTCACTGTGTACGCTGTGGACGGGGCTTTTCA	612
Db	369	CAAGGTACTTAAAGCATATAATCAGACTCACTGTGTACGCTGTGGACGGGGCTTTTCA	428
QY	613	TCCAAATTCGACCTACATATGAATTTGGAATCGACATCATCTGTAGGACAAATATTTTAAAGCTGA	672
Db	429	TCCAAATTCGACCTACATATGAATTTGGAATCGACATCATCTGTAGGACAAATATTTTAAAGCTGA	488
QY	673	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	732
Db	489	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	548
QY	733	ATCCCTTTTAAATAGATGGGAATTTACTCTGGAACCTGAGCTGATTTTATGGGGGAGA	792
Db	549	ATCCCTTTTAAATAGATGGGAATTTACTCTGGAACCTGAGCTGATTTTATGGGGGAGA	608
QY	793	CTTTGTCTATCTTCCGAACCTTTGGGCACCAACCACCAATCAGACAGACAGCATGATTC	852
Db	609	CTTTGTCTATCTTCCGAACCTTTGGGCACCAACCACCAATCAGACAGACAGCATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATGAGTCCCACTCATCTCAGAGAGTGACAAATCC	912
Db	669	CAGGTGGCTCAATGATCCAAAGTTCAATGAGTCCCACTCATCTCAGAGAGTGACAAATCC	728
QY	913	TGAAGATGACAAAGTATACCTTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTGG	972
Db	729	TGAAGATGACAAAGTATACCTTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTGG	788
QY	973	AAAAGCTACTCAGCTAGAAATAGTCCAGATATGCAAGAATGACTTTGGAGGGCACAGAG	1032
Db	789	AAAAGCTACTCAGCTAGAAATAGTCCAGATATGCAAGAATGACTTTGGAGGGCACAGAG	848
QY	1033	TCGTGGAATAAATGACAAACATTCCTCAAGCTCGCTGTGATTTGCTCAGTCCCAAGGTCC	1092
Db	849	TCGTGGAATAAATGACAAACATTCCTCAAGCTCGCTGTGATTTGCTCAGTCCCAAGGTCC	908
QY	1093	AAATGGCATTTGACCTCAATTTTGAATGAACTGCAAGGATGATTCCTTAATGAACTTTAAAGA	1152
Db	909	AAATGGCATTTGACCTCAATTTTGAATGAACTGCAAGGATGATTCCTTAATGAACTTTAAAGA	968
QY	1153	TCTTAAATATCCAGTTGTATATGGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATC	1212
Db	969	TCTTAAATATCCAGTTGTATATGGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATC	1028
QY	1213	AGCGGTGTGTATGATAGCATGAGTGTGAGAGAGGGTGTCTCTGGTCCATATATGCCCA	1272
Db	1029	AGCGGTGTGTATGATAGCATGAGTGTGAGAGAGGGTGTCTCTGGTCCATATATGCCCA	1088

QY	1273	CAGGATGGAACCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCC	1332
Db	1089	CAGGATGGAACCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCC	1148
QY	1333	AGGAACCTTTGTCCAGCAAAAACATTTTGGTGGTTTGAAGTCTTCAAAAGGACCTTCTGATGA	1392
Db	1149	AGGAACCTTTGTCCAGCAAAAACATTTTGGTGGTTTGAAGTCTTCAAAAGGACCTTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAAGTATCCAGGCCATGTACAAATTCAGTGTTCCTATGACAA	1452
Db	1209	TGTTATAACCTTTGCAAGAAGTATCCAGGCCATGTACAAATTCAGTGTTCCTATGACAA	1268
QY	1453	TCGCCCAATAGTGTACAAACGGATGTAAATTTATCAATTTACAAATTTGCTGTAGACCG	1512
Db	1269	TCGCCCAATAGTGTACAAACGGATGTAAATTTATCAATTTACAAATTTGCTGTAGACCG	1328
QY	1513	AGTGATGACAGAAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGT	1572
Db	1329	AGTGATGACAGAAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGT	1388
QY	1573	TCTTAAAGTAGTTTCAATTTCTTAAGGACACTTGGTATGATTTAGAAAGGTTTCTGCTGA	1632
Db	1389	TCTTAAAGTAGTTTCAATTTCTTAAGGACACTTGGTATGATTTAGAAAGGTTTCTGCTGA	1448
QY	1633	AGAAATGACAGTTTTTTCGGGAACCGACTGTATTTTTCAGCAATGGAGCTTTCCACTAAGCA	1692
Db	1449	AGAAATGACAGTTTTTTCGGGAACCGACTGTATTTTTCAGCAATGGAGCTTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATATTTGTTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGA	1752
Db	1509	GCAACAACTATATATTTGTTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGA	1568
QY	1753	TATTTACGGGAAGCGTGTGTGAGTGTTCCTCCACATGCAAGAGAGCCCAAGACAGATAT	1812
Db	1569	TATTTACGGGAAGCGTGTGTGAGTGTTCCTCCACATGCAAGAGAGCCCAAGACAGATAT	1628
QY	1813	TGGTTCTCATGTTCTCGCTATTTTCCACATGCAAGAGAGCCCAAGACAGATAT	1872
Db	1629	TGGTTCTCATGTTCTCGCTATTTTCCACATGCAAGAGAGCCCAAGACAGATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCTGTTTTCAGACTTACACCATGATATTAATCACCATGGCA	1932
Db	1689	AAGAAATGGAGACCCACTGACTCTGTTTTCAGACTTACACCATGATATTAATCACCATGGCA	1748
QY	1933	CAGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGCACATTTTGGAAATGCAG	1992
Db	1749	CAGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGCACATTTTGGAAATGCAG	1808
QY	1993	TCCGAAGTCGAGAGAGCGCTGTTCTATTGGCAATTTCCAGAGGCGCAATGAAGAGCGAA	2052
Db	1809	TCCGAAGTCGAGAGAGCGCTGTTCTATTGGCAATTTCCAGAGGCGCAATGAAGAGCGAA	1868
QY	2053	AGAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTTCTGCTACGTAG	2112
Db	1869	AGAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTTCTGCTACGTAG	1928
QY	2113	TCTAACACAGAGAGATTCAGGCAATTAATCTCTGCCATCGGTGGAGACATGGGTTATACA	2172
Db	1929	TCTAACACAGAGAGATTCAGGCAATTAATCTCTGCCATCGGTGGAGACATGGGTTATACA	1988
QY	2173	AACCTCTTAAAGGTAAACCTTGAAGTCAATTTGACACAGAGATTTGGAAAGAACTTCTTCA	2232
Db	1989	AACCTCTTAAAGGTAAACCTTGAAGTCAATTTGACACAGAGATTTGGAAAGAACTTCTTCA	2048
QY	2233	TAAAGATGATGAGATGGCTCTTAAGA CAAAGAAATGTCCAATAGCATGACACCTAG	2292
Db	2049	TAAAGATGATGAGATGGCTCTTAAGA CAAAGAAATGTCCAATAGCATGACACCTAG	2108
QY	2293	CCAGAGAGTCTGGTACAGAGACTTCATCGAGCTCATCAACCCCAATCTCAACAGCAT	2352
Db	2109	CCAGAGAGTCTGGTACAGAGACTTCATCGAGCTCATCAACCCCAATCTCAACAGCAT	2168

QY 2353 GGATGAGTCTCTGTGACAAAGTTTGGAAAGGACCGGAAACAAACGTCGGCAAGGCCAGG 2412
DB 2169 GGATGAGTCTCTGTGACAAAGTTTGGAAAGGACCGGAAACAAACGTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCCGAGGACAGTAACAAATGGAGGACCTTACAGAAATAAGAAAGGTAGAAA 2472
DB 2229 ACATACCCCGAGGACAGTAACAAATGGAGGACCTTACAGAAATAAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGCAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCANTTACCTCTAGA 2532
DB 2289 CAGGAGGACCCAGCAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCANTTACCTCTAGA 2348
QY 2533 AACCTCAACAGTAGAACAATTTGCGCTAGACAAATGCTGAAACAAATGCAATATACAT 2592
DB 2349 AACCTCAACAGTAGAACAATTTGCGCTAGACAAATGCTGAAACAAATGCAATATACAT 2408
QY 2593 GAACTTTTTCATGCAATTTGAGATGTTTCAATGTTTCAATGTTGGAATTTCAAGCTGAGTTCCA 2652
DB 2409 GAACTTTTTCATGCAATTTGAGATGTTTCAATGTTTCAATGTTGGAATTTCAAGCTGAGTTCCA 2468
QY 2653 CCAATTATAAATTAATCCATGAGTAACCTTCTTAATAGGCTTTTTTTTCC 2702
DB 2469 CCAATTATAAATTAATCCATGAGTAACCTTCTTAATAGGCTTTTTTTTCC 2518

RESULT 2
US-08-835-268-53
; Sequence 53, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mathes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS

LOCATION: 16...2331
US-08-835-268-53
Query Match 92.6%; Score 2508.4; DB 1; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 193 CTGACGATGGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 252
DB 9 CTGACGATGGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 68
QY 253 AAGACCAACTATCAGATGGGAGAAACAATGTGCAAGGCTGAAATTTATCTTACAAAGA 312
DB 69 AAGACCAACTATCAGATGGGAGAAACAATGTGCAAGGCTGAAATTTATCTTACAAAGA 128
QY 313 AATGTTGGAATCCCAACAATGTGATCCTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 372
DB 129 AATGTTGGAATCCCAACAATGTGATCCTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 432
DB 189 TACCTTCCTTTTGGATGAGGAAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 248
QY 433 TTCATTGACCTGTTTAATATCAAGGATTTTCAAAAGATTTGCTGGCCAGTATCTTACAC 492
DB 249 TTCATTGACCTGTTTAATATCAAGGATTTTCAAAAGATTTGCTGGCCAGTATCTTACAC 308
QY 493 CAGAAAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTTAATTTTCAAT 552
DB 309 CAGAAAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTTAATTTTCAAT 368
QY 553 CAAGGTACTTAAAGGCATATAATCAAGACTCCTTTGACGCTGTGGAACGGGGGCTTTTCA 612
DB 369 CAAGGTACTTAAAGGCATATAATCAAGACTCCTTTGACGCTGTGGAACGGGGGCTTTTCA 428
QY 613 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 672
DB 429 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 488
QY 673 GAACTCACATTTTGAACACGGCCGTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGACG 732
DB 489 GAACTCACATTTTGAACACGGCCGTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGACG 548
QY 733 ATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA 792
DB 549 ATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTTGGGACACCAACCCAAATCAGGACAGAGCATGATTC 852
DB 609 CTTTGTCTATCTTCGAACTCTTTGGGACACCAACCCAAATCAGGACAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTCTCCGTGAAAATCAATAGATGAGAGAACTCTGCG 972
DB 729 TGAAGATGACAAAGTATATCTTTCTCCGTGAAAATCAATAGATGAGAGAACTCTGCG 788
QY 973 AAAAGCTACTCACGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTGAGGGGACAGAAAG 1032
DB 789 AAAAGCTACTCACGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTGAGGGGACAGAAAG 848
QY 1033 TCTGTGAAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC 1092
DB 849 TCTGTGAAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC 908
QY 1093 AAATGGCATTCACACTCATTTTGTATGAACTCGAGGATGTTTCTTAATGAATCTTTAAAGA 1152
DB 909 AAATGGCATTCACACTCATTTTGTATGAACTCGAGGATGTTTCTTAATGAATCTTTAAAGA 968
QY 1153 TCCTAAAAATCCAGTTGTATATGAGTGTTPAAGCTTCCAGTAACTTTTCAAGGGATC 1212

Db 969 TCCTAAAAATCCAGTTGTATATGAGTGTGTTACGACTTCCAGTAAACATTTTCAAGGATC 1028
Qy 1213 AGCCGTTGTATGATAGATAGTATGAGAGGTTCTTGTGTCATATGSCCA 1272
Db 1029 AGCCGTTGTATGATAGATAGTATGAGAGGTTCTTGTGTCATATGSCCA 1088
Qy 1273 CAGGATGGAACCAACTATCAATGGTGTCTTATCAAGGAAGTCCCTTATCCAGGCC 1332
Db 1089 CAGGATGGAACCAACTATCAATGGTGTCTTATCAAGGAAGTCCCTTATCCAGGCC 1148
Qy 1333 AGAACTGTCCAGCAAAACATTTGGTGTGTTGACTCTCAAGGAACCTTCTGATGA 1392
Db 1149 AGAACTGTCCAGCAAAACATTTGGTGTGTTGACTCTCAAGGAACCTTCTGATGA 1208
Qy 1393 TGTATAAATCTTGGCAAGAGTATCCAGCATGTATCAATCCAGTGTCTTCTATGAACA 1452
Db 1209 TGTATAAATCTTGGCAAGAGTATCCAGCATGTATCAATCCAGTGTCTTCTATGAACA 1268
Qy 1453 TGCCCCAATAGTATCAAAACCGATGTAAATTTATCAAAATTTGTGTAGACCG 1512
Db 1269 TGCCCCAATAGTATCAAAACCGATGTAAATTTATCAAAATTTGTGTAGACCG 1328
Qy 1513 AGTGGATGCAAGATGGAAGTATGATGTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGGATGCAAGATGGAAGTATGATGTATGTTTATCGGAACAGATGTTGGACCGT 1388
Qy 1573 TCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1632
Db 1389 TCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1448
Qy 1633 AGAAATGACAGATTTTTCGGGAACCGACTGTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1692
Db 1449 AGAAATGACAGATTTTTCGGGAACCGACTGTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1508
Qy 1693 GCAACAACTATATATTTGGTTCAACCGCTGGGTTGCCAGCTCCCTTTACCGGTGA 1752
Db 1509 GCAACAACTATATATTTGGTTCAACCGCTGGGTTGCCAGCTCCCTTTACCGGTGA 1568
Qy 1753 TATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCGCGGAGACCTTACTGTGTTGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCGCGGAGACCTTACTGTGTTGGA 1628
Qy 1813 TGCTTCTGATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1872
Db 1629 TGCTTCTGATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1688
Qy 1873 AGAAATGAGACCCACTGACTCACTGTTTCCAGACTTACACCATGATAATCAGCATGGCCA 1932
Db 1689 AGAAATGAGACCCACTGACTCACTGTTTCCAGACTTACACCATGATAATCAGCATGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGATCATCTATGTTGATAGATAGTACACATTTTTCGATGCGAG 1992
Db 1749 CAGCCCTGAAGAGAGATCATCTATGTTGATAGATAGTACACATTTTTCGATGCGAG 1808
Qy 1993 TCCGAAGTCCGAGAGAGCGTGTCTATTTGGCAATTCAGAGCGCAATGAAGAGCGAAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCGTGTCTATTTGGCAATTCAGAGCGCAATGAAGAGCGAAA 1868
Qy 2053 AGAAGATGATAGTGGATGATCATATCATAGGACAGATCAAGGCTTCTGCTAGTAG 2112
Db 1869 AGAAGATGATAGTGGATGATCATATCATAGGACAGATCAAGGCTTCTGCTAGTAG 1928
Qy 2113 TCTACACAGAGGATTCAGGCAATTTACCTCTGCGATGCGGTGGACATGGTTCTATCA 2172
Db 1929 TCTACACAGAGGATTCAGGCAATTTACCTCTGCGATGCGGTGGACATGGTTCTATCA 1988
Qy 2173 AACTCTTCTTAAGGTAACTTGAAGTCTATGACACAGAGCATTTTGGAAAGAACTTCTCA 2232
Db 1989 AACTCTTCTTAAGGTAACTTGAAGTCTATGACACAGAGCATTTTGGAAAGAACTTCTCA 2048
Qy 2233 TAAAGATGATAGTGGATGCTTAAAGACCAAGAAATGTCCAATGATGACACCTAG 2292
Db 2049 TAAAGATGATAGTGGATGCTTAAAGACCAAGAAATGTCCAATGATGACACCTAG 2108

Qy 2293 CCAGAAGTCTGTTACAGAGACTTCAATGAGCTCATCAACACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAAGTCTGTTACAGAGACTTCAATGAGCTCATCAACACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTGTGAAACAAAGTTTGAAGGACCGAAACAAACGTCGGCAAGGCGAGG 2412
Db 2169 GGATGAGTCTGTGAAACAAAGTTTGAAGGACCGAAACAAACGTCGGCAAGGCGAGG 2228
Qy 2413 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGACCCAGCAATTTGAGAGGCGCCAGGAGTGTCTGAGCTGCAATTTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGCAATTTGAGAGGCGCCAGGAGTGTCTGAGCTGCAATTTACCTCTAGA 2348
Qy 2533 AACCTCAACAAAGTAGAATACTTGCCTAGACAATACTGAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAACAAAGTAGAATACTTGCCTAGACAATACTGAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGTCATATGTCATGTTTACATGTTGGGAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGTCATATGTCATGTTTACATGTTGGGAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATAAATCCATGAGTAATTTCTTAATAGGCTTTTTCCTTCC 2702
Db 2469 CCAATTATAAATAAATCCATGAGTAATTTCTTAATAGGCTTTTTCCTTCC 2518

RESULT 3

US-09-060-692-53
; Sequence 53, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-09-060-692-53

Query Match 92.6%; Score 2508.4; DB 2; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGCTGGTAACTAGGATGTCTGTCTTTCTCGGAGATTAATCTACAGC	252
Db	9	CTGCAGCATGGCTGGTAACTAGGATGTCTGTCTTTCTCGGAGATTAATCTACAGC	68
QY	253	AAGAGCAAACTACAGATGGGAAGAAACAATGTGCAAGGCTGAAATATCTTACAAAGA	312
Db	69	AGAGCAAACTACAGATGGGAAGAAACAATGTGCAAGGCTGAAATATCTTACAAAGA	128
QY	313	AATGTTGGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	372
Db	129	AATGTTGGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	188
QY	373	TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAAAGGATCAATATT	432
Db	189	TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAAAGGATCAATATT	248
QY	433	TTCAATTCGACCTGGTTAAATATCAAGGATTTCAAAGATTTGTCGCAAGATCTTACAC	492
Db	249	TTCAATTCGACCTGGTTAAATATCAAGGATTTCAAAGATTTGTCGCAAGATCTTACAC	308
QY	493	CAGAAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGATTTGCTAAATTCAT	552
Db	309	CAGAAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGATTTGCTAAATTCAT	368
QY	553	CAAGTACTTAAGGATATAATCAGACTCATTGTGACGCTGTGGAACGGGGCTTTTCA	612
Db	369	CAAGTACTTAAGGATATAATCAGACTCATTGTGACGCTGTGGAACGGGGCTTTTCA	428
QY	613	TCCAAATTCGACCTCAATTTGAATTTGGACATCATCTGAGGACATATTTTAACTGGA	672
Db	429	TCCAAATTCGACCTCAATTTGAATTTGGACATCATCTGAGGACATATTTTAACTGGA	488
QY	673	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTCAAGCTGCTACAGC	732
Db	489	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTCAAGCTGCTACAGC	548
QY	733	ATCCCTTTTAATAGATGGGAATTAATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	792
Db	549	ATCCCTTTTAATAGATGGGAATTAATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTGCTATCTTCGAACTCTTGGGCAACCACTGAGATGATGATGATGATGATGATGATG	852
Db	609	CTTTGCTATCTTCGAACTCTTGGGCAACCACTGAGATGATGATGATGATGATGATGATG	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTCTCAGAGAGTGACAATCC	912
Db	669	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGAGAACACTCTGG	972
Db	729	TGAAGATGACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGAGAACACTCTGG	788
QY	973	AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGATGATGATGATGATGATGATG	1032
Db	789	AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGATGATGATGATGATGATGATG	848
QY	1033	TCTGTGTAATAATGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC	1092
Db	849	TCTGTGTAATAATGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC	908
QY	1093	AAATGGCATTTGACATCTATTTTGTGAACTGCGAGATGTATTTCTTAATGAATTTAAAGA	1152

Db	909	AAATGGCATTTGACACTCAATTTTGATGAACCTCAGAGATGTATTTCTTAATGAATTTAAAGA	968
QY	1153	TCCTAAAAATCCAGTTGTATATGAGTGTATACGACTTCCAGTAACTATTTTCAAGGGATC	1212
Db	969	TCCTAAAAATCCAGTTGTATATGAGTGTATACGACTTCCAGTAACTATTTTCAAGGGATC	1028
QY	1213	AGCCGTGTATGTATATAGCATGATGTGAGAAAGGTGTCTTGTGTCATATGATGCCA	1272
Db	1029	AGCCGTGTATGTATATAGCATGATGTGAGAAAGGTGTCTTGTGTCATATGATGCCA	1088
QY	1273	CAGGATGGACCCCACTATCAATGGTGCCTTATCAGGAAGAGTCCCTTATCCAGGCC	1332
Db	1089	CAGGATGGACCCCACTATCAATGGTGCCTTATCAGGAAGAGTCCCTTATCCAGGCC	1148
QY	1333	AGGAATCTTGTCCAGCAAAAATTTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGA	1392
Db	1149	AGGAATCTTGTCCAGCAAAAATTTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTCTTCTATGAACAA	1452
Db	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTCTTCTATGAACAA	1268
QY	1453	TCGCCCAATAGTATCAAAACGAGTGTAAATATCAATTTACAAATTTGCTAGACCG	1512
Db	1269	TCGCCCAATAGTATCAAAACGAGTGTAAATATCAATTTACAAATTTGCTAGACCG	1328
QY	1513	AGTGGATCGAAGATGGACAGTATGATGTATGTTTATCGGAACAGATTTGGGACCGT	1572
Db	1329	AGTGGATCGAAGATGGACAGTATGATGTATGTTTATCGGAACAGATTTGGGACCGT	1388
QY	1573	TCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCCTGCTGGA	1632
Db	1389	TCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCCTGCTGGA	1448
QY	1633	AGAAATGACAGTTCCTGGGAACCGACTGCTATTTTACCAATGAGCTTTTCCACTAAGCA	1692
Db	1449	AGAAATGACAGTTCCTGGGAACCGACTGCTATTTTACCAATGAGCTTTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA	1752
Db	1509	GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA	1568
QY	1753	TATTTACGGGAAGCGTGTCTGAGTGTGCTCCCGGAGACCTTACTGTCTGGGA	1812
Db	1569	TATTTACGGGAAGCGTGTCTGAGTGTGCTCCCGGAGACCTTACTGTCTGGGA	1628
QY	1813	TGGTTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGAGCGCAAGACCAAGATAT	1872
Db	1629	TGGTTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGAGCGCAAGACCAAGATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCTGCTGAGCTTACACCATGATATCACCATGGCCA	1932
Db	1689	AAGAAATGGAGACCCACTGACTCTGCTGAGCTTACACCATGATATCACCATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGAGAGATAGTAGCACATTTTGGAAATGAG	1992
Db	1749	CAGCCCTGAAGAGAGAAATCATCTATGTTGAGAGATAGTAGCACATTTTGGAAATGAG	1808
QY	1993	TCCGAAGTTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGCGGAAATGAGAGCGAAA	2052
Db	1809	TCCGAAGTTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGCGGAAATGAGAGCGAAA	1868
QY	2053	AGAGAGATCAGAGTGGATGATCATATCAGGACAGATCAAGGCTTCTGCTACGATAG	2112
Db	1869	AGAGAGATCAGAGTGGATGATCATATCAGGACAGATCAAGGCTTCTGCTACGATAG	1928
QY	2113	TCTACAAAGAGGATTCAGGCAATTTACCTCTGCACTCGGCTGGAAACATGGTTTCATACA	2172
Db	1929	TCTACAAAGAGGATTCAGGCAATTTACCTCTGCACTCGGCTGGAAACATGGTTTCATACA	1988
QY	2173	AACTCTTTTAAGGTAAACCTTGGAAAGTCAATTTGACAGAGATTTTGGAAAGCTTCTTCA	2232

Db 1989 AACTCTTCTTAAGGTAACTCCGGAAGTCATTGACACAGAGCATTTGGAAAGAACTTTCTTCA 2048
QY 2233 TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
QY 2293 CCAGAAAGGTCTGGTACAGAGACTTTCATGCGAGCTCATCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAAAGGTCTGGTACAGAGACTTTCATGCGAGCTCATCAACCCCAATCTCAACACGAT 2168
QY 2353 GGATGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2412
Db 2169 GGATGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2228
QY 2413 ACATACCCAGGAGCAAGTAAACAAATGGAAGGACCTTACAGAAATTAAGAAAGGTAGAA 2472
Db 2229 ACATACCCAGGAGCAAGTAAACAAATGGAAGGACCTTACAGAAATTAAGAAAGGTAGAA 2288
QY 2473 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2348
QY 2533 AACCTCAACAGTAGAAGTTCGCTAGACAAATTAAGTGAAGAAACAAATGGAATATACAT 2592
Db 2349 AACCTCAACAGTAGAAGTTCGCTAGACAAATTAAGTGAAGAAACAAATGGAATATACAT 2408
QY 2593 GAATTTTTTCATGCGATTTGAGATGTTTCAATGGTGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAATTTTTTCATGCGATTTGAGATGTTTCAATGGTGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTAATAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTTTTCC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTTTTCC 2518

RESULT 4

US-08-833-391-53
; Sequence 53, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mattes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342

; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-833-391-53

Query Match 92.6%; Score 2508.4; DB 3; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 193 CTGACGATGGGTGTTAACTAGGATGTCGTCTTTCTGGGAGATTAATTCTTACAGC 252
Db 9 CTGACGATGGGTGTTAACTAGGATGTCGTCTTTCTGGGAGATTAATTCTTACAGC 68
QY 253 AAGACCAACTATCAGATGGGAGCAATGTGCCAAGGCTGAAATTTATCTTACAAGA 312
Db 69 AAGACCAACTATCAGATGGGAGCAATGTGCCAAGGCTGAAATTTATCTTACAAGA 128
QY 313 AATGTTGAAATCCCAACTGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTTGAAATCCCAACTGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAAACGAGTAGCTGTATCTTGGAGCAAAAGGATCAATAT 432
Db 189 TACCTTCCTTTTGGATGAGGAAACGAGTAGCTGTATCTTGGAGCAAAAGGATCAATAT 248
QY 433 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 492
Db 249 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGAGACATCTGAAAGAAATGTGCTAATTTCA 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGAGACATCTGAAAGAAATGTGCTAATTTCA 368
QY 553 CAAGTACTTAAGCATATAATCAAGTCACTTTGTCGCTTGGAAACGGGGCTTTTCA 612
Db 369 CAAGTACTTAAGCATATAATCAAGTCACTTTGTCGCTTGGAAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGCACTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 488
QY 673 GAACTCAATTTTGAAGACCGCTGGGAGAGTCCATATGACCCCTAAGCTGCTGACAGC 732
Db 489 GAACTCAATTTTGAAGACCGCTGGGAGAGTCCATATGACCCCTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGGCGAGA 608
QY 793 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGGACGATGATTC 852
Db 609 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGGACGATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGACAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTTGAAATCAATAGATGAGAAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTTGAAATCAATAGATGAGAAACACTCTGG 788
QY 973 AAAAGCTACTCACGCTAGATAGGTCAGATATGAAGATCACTTTGGAGGACAGAG 1032
Db 789 AAAAGCTACTCACGCTAGATAGGTCAGATATGAAGATCACTTTGGAGGACAGAG 848

REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: B94-002-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341
 TELEFAX: (415) 343-4342
 TELEX:

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 16..2331

US-09-060-610-53

Query Match 92.6%; Score 2508.4; DB 4; Length 2601;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGGCTGTTAACTAGGATTTGCTCTCTTTCTGGGAGATTTACTTACAGC	252
DB	9	CTGCAGCATGGGCTGTTAACTAGGATTTGCTCTCTTTCTGGGAGATTTACTTACAGC	68
QY	253	AAGAGCAAACTATCAAAATGGGAAGAACAAATGTCGCAAGGCTGMAATTTCTTACAAAGA	312
DB	69	AAGAGCAAACTATCAAAATGGGAAGAACAAATGTCGCAAGGCTGMAATTTCTTACAAAGA	128
QY	313	AATGTTGGAAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372
DB	129	AATGTTGGAAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCCCTTTGGATGAGAAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	432
DB	189	TACCTTCCCTTTGGATGAGAAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	248
QY	433	TTCAATTCGACCTGGTTAAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTTGAAGAATGTGCTAAATTCAT	552
DB	309	CAGAAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTTGAAGAATGTGCTAAATTCAT	368
QY	553	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGGCTTTTCA	612
DB	369	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGGCTTTTCA	428
QY	613	TTCAATTTGCACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	672
DB	429	TTCAATTTGCACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	488
QY	673	GAACTCACTATTTGAAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTGACAGC	732
DB	489	GAACTCACTATTTGAAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTGACAGC	548
QY	733	ATCCCTTTTAAATAGAGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	792
DB	549	ATCCCTTTTAAATAGAGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTTGTATCTTCCGAACCTTTGGGCACCAACCAATCAGGACAGACGACATGATTC	852
DB	609	CTTTTGTATCTTCCGAACCTTTGGGCACCAACCAATCAGGACAGACGACATGATTC	668
QY	853	CAGGTGGCTCAATGATGCCAAAGTTCAATTAAGTCCCACTCATCTCAGAGAGTGACAATCC	912
DB	669	CAGGTGGCTCAATGATGCCAAAGTTCAATTAAGTCCCACTCATCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGAATGATGGAGAACATCTGGG	972
DB	729	TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGAATGATGGAGAACATCTCTGG	788

QY	973	AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATAGCTTTTGAGGGCACAGAAG	1032
DB	789	AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATAGCTTTTGAGGGCACAGAAG	848
QY	1033	TCTGGTGAATAAATGGACAAACATTTCTCAAGCTCTGCTGATTTTCTCAGTGCCAGGTCC	1092
DB	849	TCTGGTGAATAAATGGACAAACATTTCTCAAGCTCTGCTGATTTTCTCAGTGCCAGGTCC	908
QY	1093	AAATGGCATTCACACTCATTTTGTATGAACTGCGAGATGATTTCTTAATGAATTTTAAAGA	1152
DB	909	AAATGGCATTCACACTCATTTTGTATGAACTGCGAGATGATTTCTTAATGAATTTTAAAGA	968
QY	1153	TCCTAAAAAATCCAGTTGTATGAGTGTTCAGACTTTCCAGTAACATTTTCAAGGGATC	1212
DB	969	TCCTAAAAAATCCAGTTGTATGAGTGTTCAGACTTTCCAGTAACATTTTCAAGGGATC	1028
QY	1213	AGCCGTGTGATGATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCCA	1272
DB	1029	AGCCGTGTGATGATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCCA	1088
QY	1273	CAGGATGGACCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGGCC	1332
DB	1089	CAGGATGGACCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGGCC	1148
QY	1333	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTTCAGTCTTCAAAAGGACCTTCTCTGATGA	1392
DB	1149	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTTCAGTCTTCAAAAGGACCTTCTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAAGTCACTCCAGCCATGTACAACTCAGTGTTCCTATGAACAA	1452
DB	1209	TGTTATAACCTTTGCAAGAAGTCACTCCAGCCATGTACAACTCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCAATAGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTGTAGACCG	1512
DB	1269	TCGCCAATAGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTGTAGACCG	1328
QY	1513	AGTGATGCAAGAAGTGGACAGATGATGTTATGTTATCGGAACAGATGTTGGACCGT	1572
DB	1329	AGTGATGCAAGAAGTGGACAGATGATGTTATGTTATCGGAACAGATGTTGGACCGT	1388
QY	1573	TCTTAAAGTAGTCTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGGA	1632
DB	1389	TCTTAAAGTAGTCTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAAGGTTCTGCTGGA	1448
QY	1633	AGAAATGACAGTCTTTCGGGAACCGACTGCTATTTTACAGCAATGAGCTTTTCCACTAAGCA	1692
DB	1449	AGAAATGACAGTCTTTCGGGAACCGACTGCTATTTTACAGCAATGAGCTTTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATTTGTTTCAACGGCTGGGGTTTGGCCAGCTCCCTTTTACACCGGTGGA	1752
DB	1509	GCAACAACTATATTTGTTTCAACGGCTGGGGTTTGGCCAGCTCCCTTTTACACCGGTGGA	1568
QY	1753	TATTTACGGGAAACGCTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTGGGA	1812
DB	1569	TATTTACGGGAAACGCTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTGGGA	1628
QY	1813	TGGTTCTGCTGCTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACAGATAT	1872
DB	1629	TGGTTCTGCTGCTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACAGATAT	1688
QY	1873	AAGAAATGGACCCCACTGACTCTGCTTACACTTACACCATGATATACCATGGCCA	1932
DB	1689	AAGAAATGGACCCCACTGACTCTGCTTACACTTACACCATGATATACCATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCAATTTTGGAAATGAG	1992
DB	1749	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCAATTTTGGAAATGAG	1808
QY	1993	TCCGAAGTCCGAGAGAGGGCTGCTTATTTGCAATTTCCGAGGCGAATAGAGCGGAA	2052
DB	1809	TCCGAAGTCCGAGAGAGGGCTGCTTATTTGCAATTTCCGAGGCGAATAGAGCGGAA	1868

QY 2053 AGAAGAGATCAGATCGATGATCATATCATCAGGACAGATCAAGGCTTCTCTAGCTAG 2112
Db 1869 AGAAGAGATCAGATCGATGATCATATCATCAGGACAGATCAAGGCTTCTCTAGCTAG 1928
QY 2113 TCTACAAACAGAGATTCAAGGCAATTAATCTCTGCGATCGGTGGAACATGGGTTCATACA 2172
Db 1929 TCTACAAACAGAGATTCAAGGCAATTAATCTCTGCGATCGGTGGAACATGGGTTCATACA 1988
QY 2173 AACTCTTCTTAAGGTAACCTGGAAGTCAATGACACAGAGCAATTTGGAAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTGGAAGTCAATGACACAGAGCAATTTGGAAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGGATGAGTCTTAAGACCAAGAAATGTCATAGATGACACCTAG 2292
Db 2049 TAAAGATGATGAGGATGAGTCTTAAGACCAAGAAATGTCATAGATGACACCTAG 2108
QY 2293 CCAGAGAGTCTGGTACAGAGACTTCAATGAGCTCAACACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAGAGTCTGGTACAGAGACTTCAATGAGCTCAACACCCCAATCTCAACACGAT 2168
QY 2353 GGATGAGTCTGTGAACAAGTTTGGAAAGGACCGAAACAAACGTCGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAAGTTTGGAAAGGACCGAAACAAACGTCGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGAAACAGTAACAATGGAAGCACTTACAGAAATAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGAAACAGTAACAATGGAAGCACTTACAGAAATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGCAATTTGAGAGGACCCAGGAGTCTGAGCTGCAATTAACCTTACA 2532
Db 2289 CAGGAGGACCCAGCAATTTGAGAGGACCCAGGAGTCTGAGCTGCAATTAACCTTACA 2348
QY 2533 AACCTCAAAACAAGTAGAACTTGCCTAGACAATACTGGAAACCAAAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAAGTAGAACTTGCCTAGACAATACTGGAAACCAAAATGCAATATACAT 2408
QY 2593 GAACCTTTTTCATGCAATTAATGAGTGTTCATGTTTCAATGTTGGAAATTCAGCTGAGTTCA 2652
Db 2409 GAACCTTTTTCATGCAATTAATGAGTGTTCATGTTTCAATGTTGGAAATTCAGCTGAGTTCA 2468
QY 2653 CCAATTTATAATTAATCCATCAGTAACCTTCCCTAATAGGCTTTTTC 2702
Db 2469 CCAATTTATAATTAATCCATCAGTAACCTTCCCTAATAGGCTTTTTC 2518

RESULT 6

PCT-US94-10151A-53
; Sequence 53, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR HOERBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10151A
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; PCT-US94-10151A-53

Query Match

Best Local Similarity 92.6%; Score 2508.4; DB 5; Length 2601;
Matches 2509; Conservative 0; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGACGATCGGCTGGTTAACTAGGATTTGTCTGTTTCTTTCTGGGAGTATTACTTTACAGC 252
Db 9 CTGACGATCGGCTGGTTAACTAGGATTTGTCTGTTTCTTTCTGGGAGTATTACTTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCGCAAGGCTGAAAATTTATCCTACAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCGCAAGGCTGAAAATTTATCCTACAAAGA 128
QY 313 AATGTGGAATCCCAACCAATGTGATCACTTTCAATGGGCTTGGCCCAACAGCTCCAGTTATCA 372
Db 129 AATGTGGAATCCCAACCAATGTGATCACTTTCAATGGGCTTGGCCCAACAGCTCCAGTTATCA 188
QY 373 TACCTTCTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT 432
Db 189 TACCTTCTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT 248
QY 433 TTCAATGCACTGTTTAAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 492
Db 249 TTCAATGCACTGTTTAAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGCTGGAAAGAGACATCTGAAAGAAATGCTGAATTTTCAT 552
Db 309 CAGAAGAGATGAATGCAAGTGGCTGGAAAGAGACATCTGAAAGAAATGCTGAATTTTCAT 368
QY 553 CAAGGTACTTAAAGGCATATATATCAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAAGGCATATATATCAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTTACATTTGAAATTTGGAATTTGGAATTTGGAATTTTAAAGCTGGA 672
Db 429 TCCAAATTTGCACTTACATTTGAAATTTGGAATTTGGAATTTGGAATTTTAAAGCTGGA 488
QY 673 GAACCTCACTTTTGAACCGGCTGGGAAGAGTCCATATGACCTTAAGCTGTCAGCAGC 732
Db 489 GAACCTCACTTTTGAACCGGCTGGGAAGAGTCCATATGACCTTAAGCTGTCAGCAGC 548
QY 733 ATCCCTTTTAAATAGATGGAATTTATATCTCTGGAATCTGCACTGATTTTATGGGGCCAGA 792
Db 549 ATCCCTTTTAAATAGATGGAATTTATATCTCTGGAATCTGCACTGATTTTATGGGGCCAGA 608
QY 793 CTTTGTCTATCTTCGGAATCTTTGGGCAACCCCAATCAGGACAGAGCAATGATTC 852
Db 609 CTTTGTCTATCTTCGGAATCTTTGGGCAACCCCAATCAGGACAGAGCAATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTAGTGGCCCTCATCTCAGAGAGTGCATATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTCATTAGTGGCCCTCATCTCAGAGAGTGCATATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGGTCAATATGATCAAGAAATGATTTTGGAGGCAAG 1032

TELECOMMUNICATION INFORMATION:

```

/ INFORMATION INFORMATION
/ TELEPHONE: 215-568-3100
/ TELEFAX: 215-568-3100
/ TELEFAX: 215-568-3439
/ INFORMATION FOR SEQ ID NO: 1
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1481 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: both
/ MOLECULE TYPE: CDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 50..1480
/ US-08-136-922-1

```

Query Match 52.2%; Score 1415.2; DB 1; Length 1481;
Best Local Similarity 99.4%; Pred. NO. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	899	GAGAGTGACAATCCTGAAGATGACAAAGTATAC	TTTTTCTTCCGTGAAATGCAATAGAT	958
DB	53	GAACATGACAATCCTGAAGATGACAAAGTATAC	TTTTTCTTCCGTGAAATGCAATAGAT	112
QY	959	GGAGACACTCTGGAAAAGCTACTCAGCTAGAA	TAGGTGAGATATGCAAGATGAC	1018
DB	113	GGAGACACTCTGGAAAAGCTACTCAGCTAGAA	TAGGTGAGATATGCAAGATGAC	172
QY	1019	GGAGGGCACAGAAGTCTGGTGAATAAATGACA	CAACTTCTCAAAGCTCGTGTGATTTGC	1078
DB	173	GGAGGGCACAGAAGTCTGGTGAATAAATGACA	CAACTTCTCAAAGCTCGTGTGATTTGC	232
QY	1079	TCAGTCCAGGTCCTCAATGGCAATTTGACACT	CATTTCGATGAATCGCAGGATGTATTCCTTA	1138
DB	233	TCAGTCCAGGTCCTCAATGGCAATTTGACACT	CATTTCGATGAATCGCAGGATGTATTCCTTA	292
QY	1139	ATGAACHTTTAAAGATCCTAAATAATCCAGTT	GTATATGAGATGTTTACGACTTCCAGTAAAC	1198
DB	293	ATGAACHTTTAAAGATCCTAAATAATCCAGTT	GTATATGAGATGTTTACGACTTCCAGTAAAC	352
QY	1199	ATTTCCTCAAGGGATCAGCCGCTGTGTATCTA	TAGCATGATGATGTGAGAAAGGCTGTTCCCTT	1258
DB	353	ATTTCCTCAAGGGATCAGCCGCTGTGTATCTA	TAGCATGATGATGTGAGAAAGGCTGTTCCCTT	412
QY	1259	GGTCCATATGCCACAGGGATGGACCCAACTA	CAATGGGTGCTTATCAAGGAAAGATC	1318
DB	413	GGTCCATATGCCACAGGGATGGACCCAACTA	CAATGGGTGCTTATCAAGGAAAGATC	472
QY	1319	CCCTATCCAGCGCAGGAACTGTCCAGCAAA	CATTTCGGTGGTTTGTACTTCAAAAG	1378
DB	473	CCCTATCCAGCGCAGGAACTGTCCAGCAAA	CATTTCGGTGGTTTGTACTTCAAAAG	532
QY	1379	GACCTTCCTGATGATGTATAACCTTTGCAAG	AGATCATCCAGCCATGTACAAATCCAGTG	1438
DB	533	GACCTTCCTGATGATGTATAACCTTTGCAAG	AGATCATCCAGCCATGTACAAATCCAGTG	592
QY	1439	TTTTCTATGAACAATCGCCCAATAGTGATCA	AAACCGGATGTAAATTTACAAATTCACAA	1498
DB	593	TTTTCTATGAACAATCGCCCAATAGTGATCA	AAACCGGATGTAAATTTACAAATTCACAA	652
QY	1499	ATTGTCGTAGACCGAGTGGATGACAGAAATG	GCACAGTATGATTTATGTTTATCGGAACA	1558
DB	653	ATCGTCGTAGACCGAGTGGATGACAGAAATG	GCACAGTATGATTTATGTTTATCGGAACA	712
QY	1559	GATGTTGGGACCGTTCCTTAAAGTAGTTTCA	ATTTCTTAAGGAGACTTGGTATGATTTAGAA	1618
DB	713	GATGTTGGGACCGTTCCTTAAAGTAGTTTCA	ATTTCTTAAGGAGACTTGGTATGATTTAGAA	772
QY	1619	GAGGTTCTGCTGGAAGAAATTCACAGATTTT	TCGGGAAACCGACTGTATTTTCAGCAATGGAG	1678
DB	773	GAGGTTCTGCTGGAAGAAATTCACAGATTTT	TCGGGAAACCGACTGTATTTTCAGCAATGGAG	832
QY	1679	CTTTCCAATGAAGACCAACCACTATATATTT	TGGTTTCAACCGCTGGGGTGGCCAGCTCCCT	1738

Db	833	CTTTCACCTAAGCAGCAACAACATATATATATGGTTCAACGGCTGGGGTGTGCCAGCTCCCT	892
Qy	1739	TTACACCGGTGTGATATTTACGGGAAAGCGTGTGCTGAGTGTGGCTCGCGCCGAGACCCCT	1798
Db	893	TTACACCGGTGTGATATTTACGGGAAAGCGTGTGCTGAGTGTGGCTCGCGCCGAGACCCCT	952
Qy	1799	TACTGTCTTGGGATGGTCTGCAATGTTTCGCTATTTTTCCCACTGCAAAAGACAGCAACA	1858
Db	953	TACTGTCTTGGGATGGTCTGCAATGTTTCGCTATTTTTCCCACTGCAAAAGACAGCAACA	1012
Qy	1859	AGACGACAAGATATAGAATAGACCCCACTCACTCACTGTTTCAGTCTTACACCATGAT	1918
Db	1013	AGACGACAAGATATAGAATAGACCCCACTCACTCACTGTTTCAGTCTTACACCATGAT	1072
Qy	1919	AATCACCATGCCACAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGAATAGTAGCACA	1978
Db	1073	AATCACCATGCCACAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGAATAGTAGCACA	1132
Qy	1979	TTTTTGGATGCAGTCGGAAGTCGACAGAGAGCGCTGGTCTATTGGCAATTCACAGAGCGGA	2038
Db	1133	TTTTTGGATGCAGTCGGAAGTCGACAGAGAGCGCTGGTCTATTGGCAATTCACAGAGCGGA	1192
Qy	2039	AATGAAGAGCGAAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGAAATCAAGGC	2098
Db	1193	AATGAAGAGCGAAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGAAATCAAGGC	1252
Qy	2099	CTTCTGCTACGTAGTCTCAACAAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAA	2158
Db	1253	CTTCTGCTACGTAGTCTCAACAAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAA	1312
Qy	2159	CATGGGTTTCATAAAGCTCTTCTTAGGTTAAACCTTGGAGTCAATTGACACAGAGCATTTG	2218
Db	1313	CATGGGTTTCATAAAGCTCTTCTTAGGTTAAACCTTGGAGTCAATTGACACAGAGCATTTG	1372
Qy	2219	GAAGAACTTCTTCATAAAGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCTCAAT	2278
Db	1373	GAAGAACTTCTTCATAAAGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCTCAAT	1432
Qy	2279	AGCATGACACCTAGCCAGAGGTCCTGGTACAGAGATCTTCATGCGTCTC	2326
Db	1433	AGCATGACACCTAGCCAGAGGTCCTGGTACAGAGATCTTCATGCGCGCC	1480

RESULT 8

```

US-09-308-179B-2
; Sequence 2, Application US/09308179B
; Patent No. 6436669
; GENERAL INFORMATION:
; APPLICANT: INAGAKI, Shinobu
; APPLICANT: FURUYAMA, Tatsuo
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (I)
; FILE REFERENCE: 0020-4562P
; CURRENT APPLICATION NUMBER: US/09/308,179B
; CURRENT FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/JPS97/04111
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: JAPAN 321068/1996
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: Strandedness: Double-stranded
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: any n = a, c, g, t, unknown, or other
US-09-308-179B-2

```

Query Match		22.0%;	Score 596.4;	DB 4;	Length 2898;
Best Local Similarity		56.3%;	Pred. No. 11e-175;		
Matches 1229;		Conservative	0;	Mismatches 931;	Indels 24; Gaps 5;
QY	223	CTGTCTTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAGAATGGGAAGAACAA	282		
DB	396	CTTGTGCTCTGGGTCACCTGCTGGAACCTCTGGACCCCGAGGTCACTCCGCGAACCCCTC	455		
QY	283	TGTGCCAAGCTGAAATATCTTACAAAGAAATGTTGGAAATCCAAATGTGATCTACTTT	342		
DB	456	CTACCCGAGCTACGGCTGTGCATTAAGAACTTTTGGAACTGAATGAGGACTTCAATAT	515		
QY	343	CAATGCTTGGCCAAACAGCTCCAGTTATCATACCTTCTCTTGGATGAGGAACGGAGTAG	402		
DB	516	TCAAAGCCCTTGGATTTCTTGATCTCCATACATGCTGCTGATGATATCAAGACG	575		
QY	403	GCTGTATGTTGGAGCAAAAGGATCACATATTTTCATTCGACCTGGTTAAATATCA--AGGA	459		
DB	576	GCTCTTTTGGGAGGCAGAGACCTTGTCTATTCCTGAACTTGGACGAGTCAGTCAAGG	635		
QY	460	TTTTTCAAAGATTTGTGTGGCAGTATCTTACACCAAGAGAGATGAATGCAAGTGGGCTGG	519		
DB	636	CTACAGAGAGATATACTGGCCGAGCACAGCAGTAAAGGTAGAAGATGCAATGAAGG	695		
QY	520	AAAAGACATCTGAAGAAATGTCTAAATTTTCATCAAGGTACTTAAAGGCATATAATCAGAC	579		
DB	696	AAAGAA--CGCAATGAGTGTGCCAATTAATATCCGGGTTTTCGATCACTACACAGGAC	752		
QY	580	TCATTTGTAGCTGTGGAAACGGGGCTTTTTCATCCAAATTTGCACCTACATTTGAAATGG	639		
DB	753	ACACCTTCTGACCTGTACTGGAGCTTTTGTATCCACACTGTGCTTCATCAGAGTCGG	812		
QY	640	ACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGTGG	699		
DB	813	GCACCAATCAGAGAAACCCCTGTCTTCACTGGAGTCAACAGATCTGAGAGAGGAAGGG	872		
QY	700	GAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATATA	759		
DB	873	CAGATGTCTTTTGAACCCCACTCCTCTTGTGTCACGCTAGTTGGGAATGAGCTGTT	932		
QY	760	CTCTGAACTGCACTGATTTTATGGGGCGGAGACTTTGCTATCTTCCGAACTCTTGGGCA	819		
DB	933	TGCTGACTCTACAGTGACTATTGGGGCAGAGACTCGGCGATCTTCCGACGATGGGAA	992		
QY	820	CCACACCAATCAGACACAGAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAAT	879		
DB	993	GTTAGGCCATATTCGCACCTGAGCATGACGATGAGCGGCTCCTGAAAGAACCAAAATTTGT	1052		
QY	880	TAGTGCCACCTCATCTCAGAGAGTGACAATCTGGAAGATGACAAGTATACCTTTTCTT	939		
DB	1053	AGGTTCAATATGATTCCTGTAACAGACCGAGATGACAAATGATCTTTTCTT	1112		
QY	940	CGGTGAAATGCAATAGATGGAGAACACTCTGGAAGGCTACTCAGCTAGATAAGTCA	999		
DB	1113	TACTGAGAAGCGCTGGAGCGGAGAACCAAGCCACACGATCTACACCCGAGTGGGGC	1172		
QY	1000	GATATGCAAGATGACTTTGGAGGGCAGAGAGTCTGTGTGAATGAATGACAACTCTCT	1059		
DB	1173	GCTGTGCGTGAATGACATGGGAGGACAGAGAAATCTCTGTGTAACAAGTGGAGCACTTCT	1232		
QY	1060	CAAGCTCTGTGATTTGCTCAGTGGCCAGTCCAAATGCAATGCACTCATTTTGTATGA	1119		
DB	1233	TAAAGCGCGGTGGTTTGTCTAGTCCGGGAATGAATGAATGCAATCTTTTGAAGA	1292		
QY	1120	ACTGCAAGATGATTTCTTAATGAATTTTAAAGATCTTAAATTCAGATTTGATATGAGT	1179		
DB	1293	ACTAGAGGATGTGTTTTTACTGCGGACACAGATCTTAAGATCCAGTATATTTGGACT	1352		
QY	1180	GTTTACGATCTCCAGTAAATTTTCAAGGATCAGCGGTGTGATGATATGATGATGATGA	1239		
DB	1353	GTTTAAATATACAGCAATATATTTAGAGGCCATGCTGTATGTGTATCATGTCAAG	1412		
QY	1240	TGTGAGAAAGGTGTTTCTTGTGTCATATGCCCCACAGGATGGAACCAACTATCAATGGGT	1299		
DB	1413	TATCGGGAAGCCTTTATGCCCCATATGCTCATAAAGAGGCCCTGAATACCACTGGTC	1472		
QY	1300	GCTTATCAAGAGAGTCCCTTATCCAGGCGCAGAACTTGTGCCAGCAAAACATTTGG	1359		
DB	1473	ACTATATGAAGGAAGTCCCTACCCAAAGGCTGTTCTGTGCGCAGCAAAAGTAAACGG	1532		
QY	1360	TGG---TTTTGACTCTACAAAGGACCTTCTGTATGATGTTTAACTTTGCAAGAGTCA	1416		
DB	1533	AGGCAAGTATGGAACCAACCAAGATTACCCGATGACGCCATCCGGTTCGCAAGATGCA	1592		
QY	1417	TCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATGCCCAATAGTAGATCAAAACGGA	1476		
DB	1593	TCCTTAATGTATCAGCCCATTAACCTGTTCTATTAATAAACAATATCTGTTAAAAACAGA	1652		
QY	1477	TGTAAATTTCAATTTTACAAATTTGTCGTAGACCGAGTGGATGAGAAATGGAAGTGA	1536		
DB	1653	TGGAATAATACAACCTGAGGCAACTTGGCGTGGATCGGGTGGAGCGGAGGATGGCCAGTA	1712		
QY	1537	TGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTCATATTCCTAA	1596		
DB	1713	TGACGTCTTATTTTATTTGGGACAGACACAGGAATTTGTCTGAAAGTAAATCAACAATTTACAA	1772		
QY	1597	GGAGACTTTGGTATGATTTAGAAAGAGGTTCTGCTGGAAGAAATGACAGTTTTTTTCGGGAACC	1656		
DB	1773	CCNAGAAACAGATGGATGGAGGAAGTCATTTCTAGAGGAATCTTCAATATTTCAAGGATCC	1832		
QY	1657	GACTGTATTTTCAAGATGGAGCTTTTCCATPAAGCAGCAACAATATATATTTGTTTCAAC	1716		
DB	1833	AGCCCTATCATTTTCTATGGAATTTCTTCAAGAGACACACAGCTTTTACATTTGGATCAGC	1892		
QY	1717	GGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTACGGGAAGCGTGTGCTGA	1776		
DB	1893	CTCTGCTGTGCAACAAGTCAGATTCATCTGCGACATGTATGCGAGTGTCTGTGCTGA	1952		
QY	1777	GTGTGCTCTCCCGAGACCTTACTGTCTTGGGATGGTTCTGATGTTCTCGCTATTT	1836		
DB	1953	CTGTGCTCTGCTCGAGACCGTACTGTGCTGGATGGCATATCTCTCTCCAGTACTA	2012		
QY	1837	TCCAC-----TGCAAGAGACGCAAGACGACAAGATATAAGAAATGGAAGACC	1887		
DB	2013	CCCAACAGGTGCAACGCAAGAGAGGTTCCGAGGAGGACGTTCCGCATGGCAAGCC	2072		
QY	1888	ACTGACTCAGTCTTACAGACTTACACCATGATAATCACCATGCCCCACGCTTGAAGAG	1947		
DB	2073	CGCCCAACAGTCTTTTGGACAGCAATTTGTTGGAGACGCTTGGACAGGACTGAAGAGAG	2132		
QY	1948	AATCATCTATGCTGTAGAGAAATAGTAGCACATTTTGGAAATCCAGTCCGAAAGTCCGAGAG	2007		
DB	2133	GCTGGCTTATGGCATAGAGCAACAGTACTCTGTTGGAATGCAACCCCGCATCACTACA	2192		
QY	2008	AGCGTGGTCTATTGGCAATTTCCAGAGCGGAAATGAAGAGCGGAAAGAGAGATCAGAGT	2067		
DB	2193	AGCAAAAGTCACTGCTTTGTACAGAGGGACGCGACGTAAGAAAGAAAGAGGTGAAGAC	2252		
QY	2068	GGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAGTCTACAAACAGAGA	2127		
DB	2253	GGATGACAGAGTGTCAAGATGGAAGTGGGGCTTGTCTCTCTCAGAGTACGCAAGTCTAGA	2312		
QY	2128	TTCAGGCAATTTACCTCTGCCATGCGGTGGAACATGGGTTTCATACAAACTCTTTTAAAGT	2187		
DB	2313	TGAGGGACCTTATTTTGGCAGACAGTAGAACAATTTTGTCTCCATCTGTGCGTAAAT	2372		
QY	2188	AACCTTGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGATGATGAGT	2246		
DB	2373	CACCTTGGAGTGTGCGAAGAGCAATAAGTGGAGGGCATGTTTTCATAAGGACCATGAAGA	2432		
QY	2247	-----GAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACCTAGCCAGAGGT	2301		
DB	2433	GGAAAGACATCAAGATGCGCCCTGCTCCCTTTAAGCGGTATGTCTAGGGGACAAACACC	2492		
QY	2302	CTGTACAGAGACTTTCATGCACTCATCAACCAACCCCAATCTCAACACGATGGATGAGTT	2361		

94

464 CAAAAGATTGTGGCCAGTATCTTACACGAGAGAGATGAATGCAAGTGGGCTGAAAA 523
154 CRAAGAGTACTGTGGAGTGCAGATGCTGACAGGACGACGAGTCTCAAGGCAAG 213
524 GACATCCTGAAGAAATGTGCTAATTTATCAACAGGTAATTAAGGCAATTAATCAGACTAC 583
214 GACCCAAAGCGTGACTGTCAAAACTACATCAAGATCTCTGCGCACTCAACAGCAGCCAC 273
584 TTGTAGCCCTGTGGAAGCGGGCTTTTCATCCAAATTTGACCTACATTCGAAATGGACAT 643
274 CTGCTCACCTGTGGCAGCGCGCTTACAGCCCTGTGCTTACATTCACATAGGAGC 333
644 CATCTCTGAGGACAATATTTTAAAGCTGGGAACCTCACATTTTGAAAAACGCGCTGGGAAG 703
334 TTATCTTAGCCCA---AGATGAGCGCGTAAATGTCTATCTGGAGATGCGAAGGTCA 390
704 AGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATATATCTCT 763
391 TGTCCCTTTGACCCCAACTTCAAGTCCAGCGCTCTGGTGGTGTGATGGTGAAGTGTACACT 450
764 GGAACCTGCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGAACTCTTGGGACACAC 823
451 GGAACAGTCAGTAGCTTCCAGGGAACAGCCAGCCATTTCCCGGAGCCAGAGTTCCCGC 510
824 CACCCAATCAGGACAGCAGCATGATTCAGGTGGCTCAATG---ATCCAAAGTTCAAT 880
511 CCACCAAGACTGAGAGCTCCCTCACTGGCTACAAGACCTTCTTGGCCCTCGCT 570
881 AGTGCCCACTCATCTCAGAGTGAATCTCTGAAAGATGACAAAGTATATCTTTCTTC 940
571 ACCTCCCGGAGAGCTGGGACCCCATAGGTGATGATGAATAGATCTACTTCTTCTTC 630
941 CGTGAATGCAATAGATGGAGAACACTCTGGAAGACTACTCACGCTAGATAGTCTAG 1000
631 AGCGAGACGGGCGAGAGTTTGTAGTTCTTTGAGAACACCATCGTGTCCGAGTGGCCGA 690
1001 ATATGCAAGAAATGACTTTGGAGGGCACAGAACTCTGGTGAATAAATGGACAACATTCTTC 1060
691 GTCTGTAAAGGCGATGAGGTGGAGAGCGGGTGTTCAGCAACGCTGGACCTCTTCTC 750
1061 AAGCTCTGTGATTTGCTCAGTGCCAGGT 1090
751 AAGGCTCAGCTCTGTGCTCCCGGCTGAT 780

RESULT 11

US-08-556-422A-1
; Sequence 1, Application US/08556422A
; Patent No. 6576754

GENERAL INFORMATION:

; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: HALL, Kathryn T.
; APPLICANT: SCHULTZE, Joachim J.
; APPLICANT: BOUSSIORIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-00SCP2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4157

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(2673)

US-08-556-422A-1

Query Match 3.1%; Score 83.2; DB 4; Length 4157;
Best Local Similarity 49.0%; Pred. No. 3.5e-15;

Matches 619; Conservative 0; Mismatches 568; Indels 75; Gaps 12;
QY 352 GSCCAACAGCTCCAGTTATCATACCTTCTTTTGGATGAGGAACGAGTAGGCTGTATCT 411
Db 219 CCAGACATCTTACAATCTACTAGCTTCTGCTGAGCGGAGCAAGGACACCTTGTATCT 278
QY 412 TGGCAAAAGGATCACATATTTTCAATCGACCTCGTTAATAT---CAAGGATTTTCAAAA 468
Db 279 AGTGCSCGGGAGGGGCTTTCGCTGTGAACGCACTCAACATCTCCGAGAACGACATGA 338
QY 469 GATTGTGGCCAGTATCTTACACAGAAAGAGATGAATGCAAGTGGGCTGGAAAAGACAT 528
Db 339 GGTGTATTTGGAAGTCTCAGAAGACAAAAGCAAAATGTGCAGAAAAGGGGAAATCAAA 398
QY 529 COTGAAAGAAATGTCTAATTTTCACTCAAGGTACTTAAAGGCATATAATCAGACTCACTTGA 588
Db 339 ACAGACAGATGCTCAACTACATCCGGGTGCTGAGCCACTCAGCGGCACCTTCCCTTTA 458
QY 589 CGCCTGTGGAACGGGGGCTTTTTCATCCAAATTTGCACTTACATTTGAAATTTGGACATCATCC 648
Db 459 CGTGTGCGGACCAACGCAATTCAGCGCGCTGTGACCACCTGAACCTTAAATC----- 512
QY 649 TGAGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACCGGCCCTGGGAAGAGTCC 708
Db 513 -----CTTTAAGTTTCTGGGGAATA-----TGAAGATGGCAAGGAAGATGTCC 557
QY 709 ATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATATATCTTGGAAAC 768
Db 558 CTTTGACCCAGCACAGCTACATCCGTGATGGTGTGATGGAGAACTTTATTTCCGGGAC 617
QY 769 TGCAGCTGATTTTATGGGCGAGACTTTTGTCTATCTTCCGAACTCTTGGGCAACCAACCC 828
Db 618 GTCTGATAATTTTGGGAAGTGAACCATCATCTCCGAAATCT---TCCACAGTCC 674
QY 829 AATCAGGACAGCAGCATGATTCAGGTGCTCAATGATCAAAAGTTCAATTAGTGCCA 888
Db 675 TCTGAGGACAG---AATATGCAATCCCTTGGCTGAAACGAGCTAGTTTCTGTTTGGTGA 731
QY 889 CTTCAAT-----CTCAGAGAGTGCAATCTTGAAGATGACAAAGTATATCTTTTCTT 939
Db 732 CGTATTCGAAAAAGCCAGACAGCCCGGAGCGGAGGATGACAGGGTCTACTTCTTCTT 791
QY 940 CCGTGAATAATGCAATAGATGGAGAACACTCTTGGAAAAAGCTACTCACGCTAGATAGTCA 999
Db 792 CACGAGGTGCTGTGGAGTATGATTTGTGTTCAGGGTGTGATCCCAACGATAGCAAG 851
QY 1000 GATATGCAAGAAATGACTTTTGGAGGCAAGAGTCTGTGTGAATTAATGAGCAACATTCCT 1059
Db 852 AGTGTGCAAGGGGACCCAGGGCGGCTGAGGACCTTGCAGAAAGAAATGGACCTCCTTCT 911
QY 1060 CAAAGCTCGTCTGATTTTGTCTCAGTGCAGGTCCAAATGGCAATTCACACTCAATTTGATGA 1119
Db 912 GAAAGCCGACTCATCTGCTCCCGCCAGA-----CAGCGCTTGGTCTTCAATGT 962
QY 1120 ACTGCAGGATGATTTCTTAATGAACTTTTAAAGATCTTAAAAATCCAGTTGTATGGAGT 1179
Db 963 CTGCGGATGCTCTTGTGCTCAGTCCCGGCTGAGGTGCTGTGTTCTATGCACT 1022
QY 1180 GTTACAGCTTCCAGTAACTTTTCAAGGGATACGCGGTGTGTATGTATAGATGAGTGA 1239
Db 1023 CTTCAACCCACAGCTGAACAAACGTTGGGCTGTGCGCAGTGTGCGCTTCAACCTTCCAC 1082
QY 1240 TGTGAGAGGGGTGTTCC-----TTGGTCCATATGCCCAAG---GGATGGACCCAA 1287
Db 1083 AGCCGAGAGGTCTTCTCCCAACGGGAAGTACATGAGAGCACCAAGTGGAGCAGTCCCA 1142
QY 1288 CTATCAATGGGTGCTTTATCAAGGAAGAGTCCCTTATCCAGGGCAGGAACCTTGTCCAG 1347
Db 1143 CACCAAGTGGGTGCTTATATGCCCCGTTACCAAGCCGCGCTGGAGCGTGCATCGA 1202
QY 1348 CAAACATTTGGTGGTGTGATCTCTACAAA-----GGACCTTCTGTATGTATTAAC 1401
Db 1203 CAGCGAGGACAGGGCGCCCAACTACACAGCTCCTTGAATTTGGCAGACAAGACGCTGCA 1262

Db	1673	AACAAGTCCGAACACCCAGATGGGAAAGCGATGATGTCCTGTTATAGGAACGTGATGATG	1732
Qy	1566	GGACCGTCTTAAAGTAGTTTCAATTCCTTAAGGAGACTTCGGTATGATTAGAAGAGGCTTC	1625
Db	1733	GCAAGTGATAAAGCTTGAACCTCGCCCTCTTGGTATTCATCTGATAGTAGTAGT	1792
Qy	1626	TGCTGGAAGAA 1636	
Db	1793	TTGTAATAGAA 1803	
RESULT 14			
US-09-060-692-57			
; Sequence 57, Application US/09060692			
; Patent No. 5935865			
; GENERAL INFORMATION:			
; APPLICANT: Goodman, Corey S.			
; APPLICANT: Kolodkin, Alex L.			
; APPLICANT: Matthes, David			
; APPLICANT: Bentley, David R.			
; APPLICANT: O'Connor, Timothy			
; TITLE OF INVENTION: The Semaphorin Gene Family			
; NUMBER OF SEQUENCES: 100			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP			
; STREET: 268 Bush Street, Suite 3200			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94104			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/060,692			
; FILING DATE:			
; CLASSIFICATION: 514			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/121,713			
; FILING DATE: 13-SEP-1993			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Osman, Richard A.			
; REGISTRATION NUMBER: 36,627			
; REFERENCE/DOCKET NUMBER: B94-002-1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (415)343-4341			
; TELEFAX: (415) 343-4342			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 57:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2854 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: CDNA			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 451..2640			
US-09-060-692-57			
Query Match 2.6%; Score 69.8; DB 2; Length 2854;			
Best Local Similarity 48.0%; Pred. No. 4.3e-11;			
Matches 351; Conservative 0; Mismatches 347; Indels 33; Gaps 4;			
Qy	918	ATGACAAGTACTTTTTCTTCGCGAAATGCAATAGATGAGAACACTCTGGAAAG	977
Db	1094	ATGATTTTATATCTTCTTCGAGAGACTGCTGTTAGTACATCACTCGGAAGG	1153
Qy	978	CTTACTCACGCTAGATAGGTTCAGATATGCAAGAATGACTTTTGGAGGGCAGAGAAGTCTGG	1037

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search.time 129 Seconds
(without alignments)
10838.209 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttcatgatg.....agggttttttctctaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spo1/US09774490/runat_03082003_095416_25331/app_query_fasta_1.2887
-DB=SPTEMBL_23 -QWTF=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09774490 @CGN 1 1 0 @runat_03082003_095416_25331 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTEMBL_23:*
1: sp archaea.*
2: sp bacteria.*
3: sp fungi.*
4: sp human.*
5: sp invertebrate.*
6: sp mammal.*
7: sp mhc.*
8: sp organelle.*
9: sp phage.*
10: sp plant.*
11: sp rodent.*
12: sp virus.*
13: sp vertebrate.*
14: sp unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	No.	Score	Match	Length	ID	Description
1	3609.5	73.7	774	13	Q8JIW9		Q8JIW9 xenopus lae

ALIGNMENTS

RESULT 1

ID	Q8JIW9	PRELIMINARY;	PRT;	774 AA.
AC	Q8JIW9;			
DT	01-OCT-2002 (TRENBLrel. 22, Created)			
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)			
DE	Semaphorin 3A.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Head;			
RA	Tannahill D., Nielsen J., Regan A.G.;			
RT	"Xenopus Semaphorin 3A.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY030051; AAK38166.1; -			
DR	InterPro; IPR003599; Ig.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR001005; MYB DNA binding.			
DR	InterPro; IPR003659; Plexin-like.			

Q8BH34 mus musculu
Q8BMT6 mus musculu
Q8CB71 homo sapien
Q8CDV7 homo sapien
Q8QX23 mus musculu
Q8QGU9 gallus gall
Q8G9X0 homo sapien
Q8NS98 homo sapien
Q8NCQ1 homo sapien
Q8HBR1 homo sapien
Q8BK6 mus musculu
Q8JJ29 rattus norv
Q8BJC1 mus musculu
Q8C0C4 homo sapien
Q8C0B8 homo sapien
Q8BI6 mus musculu
Q84948 mus musculu
Q8P2H9 homo sapien
Q8H2E6 homo sapien
Q8BIC3 mus musculu
Q8NFY4 homo sapien
Q8BUT0 mus musculu
Q8BXZ7 mus musculu
Q8NFY5 homo sapien
Q8BK8 mus musculu
Q8WUA9 homo sapien
Q8NFY6 homo sapien
Q8NFY3 homo sapien
Q8P249 homo sapien
Q8EQ71 mus musculu
Q8NLF1 drosophila
Q8V7Q7 drosophila
Q8P283 homo sapien
Q8GIF8 homo sapien
Q8LY36 mus musculu
Q8BXU8 mus musculu
Q8PXR8 homo sapien
Q8NFY7 homo sapien
Q8GJU9 homo sapien
O44253 drosophila
Q8V3M4 drosophila
Q8NAN9 homo sapien
Q8V7P8 drosophila
Q8R4U4 rattus norv

DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; ig; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; IG; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS00835; IG; LIKE; 1.
 DR PROSITE: PS00037; MYB; 1; 1.
 SQ SEQUENCE 774 AA; 89319 MW; 3127FD2E0CDSFA4 CRC64;

Alignment Scores:
 Pred. No.: 5,06e-298 Length: 774
 Score: 3609.50 Matches: 654
 Percent Similarity: 93.67% Conservative: 71
 Best Local Similarity: 84.50% Mismatches: 46
 Query Match: 73.72% Indels: 3
 DB: 13 Gaps: 3

US-09-774-490-1 (1-2709) x QBUIW9 (1-774)

QY 200 ATGGCGTGTAACTAGGATTGCTGCTTTCTGGGAGTATTACTTACACAGAGCA 259
 Db 1 MetGlySerLeuThrTrpIleAlaPheLeuGlyValSerLeuThrThrArgLeu 20
 QY 260 AACTATCAGAAATGGGAAGCAATGTGCCAAGCTGAAATATCTACAAAGAAATGTG 319
 Db 21 AsnCysGlnAsnValLysAsnAsnValProArgLeuArgLeuSerTyrLysGluMetVal 40
 QY 320 GAATCCAAACAACTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATACCTTC 379
 Db 41 GluSerAsnAsnLeuIleThrPheAsnGlyTyrProAsnSerSerSerTyrAsnThrPhe 60
 QY 380 CTTTGGATGAGAACGGAGTAGGTAGCTGATGTGGAGCAAGAGTACACATATTTTCATTC 439
 Db 61 ValLeuAspGluLysArgGlyArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGATTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAAGA 499
 Db 81 AsnLeuAlaAsnValLysGluPheGlnLysIleLeuTyrProValThrAsnIleArg 100
 QY 500 GATGAATGCACTGGCTGGAAGAAGACATCCTGAAAGATGTGCTAAATTTTCATCAAGGTA 559
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleValLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATAATCAGACTCATTGTACGCTGTGGACGGGGCTTTTCATCCAATT 619
 Db 121 LeuLysAlaTyrAsnHisThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProVal 140
 QY 620 TGCACCTACATTGAATTCGACATCATCTGTAG--GACAAATTTTAAAGCTGAGAAC 676
 Db 141 CysThrTyrIleAspValGlyHisTyrProGluGlnAspAsnValPheLysLeuGluAsp 160
 QY 677 TCACATTTTGAAGACGGCTGGGAAGAGTCCATATGACCCTAAGCTGTGACAGCATCC 736
 Db 161 SerPhePheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSer 180
 QY 737 CTTTAAATAGATGGAATATATCTGTGAACCTGAGCTGATTTATGGGGGAGACTTT 796
 Db 181 IleLeuIleAspGlyLysLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPhe 200
 QY 797 GCTATCTCCGAACCTCTGGGCACACACCAATCAGACAGAGCAGCATGATCCAGG 856
 Db 201 AlaIlePheArgThrLeuGlyAsnHisHisProIleArgThrGluGlnHisAspSerArg 220
 QY 857 TGGCTCAATGATCCAAATGCTATTAGTGCACCTCATCTCAGAGAGTGAATCTCTGAA 916
 Db 221 TrpLeuAsnAspProArgPheIleGlyAlaTyrLeuValProGluSerAspAsnProGlu 240
 QY 917 GATGACAAAGTATCTTTCTCGTGAATATGCAATAGATGGAGAACACTCTCGGAAA 976
 Db 241 AspAspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisThrGlyLys 260
 QY 977 GCTACTCAGCTAGAAATAGGTGAGATATGCAAGATGACTTTGGAGGGGCACAGAGTCTG 1036

Db 261 AlaThrHisAlaArgIleGlyGlnLeuCysLysAsnAspPheGlyGlyHisArgSerLeu 280
 QY 1037 GTGAATAAATGGACAAATCTCTCAAGCTCGTCTGATTGCTCAGTCCGAGGTCAAAT 1096
 Db 281 ValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsn 300
 QY 1097 GGCATTGACACTCATTTTGTGAATCGAGATGTTCTTCAATGAACCTTTAAAGATCCT 1156
 Db 301 GlyIleAspThrHisPheAspGluLeuGlnAspValValLeuMetAsnSerLysAspPro 320
 QY 1157 AAAATCCAGTTGATGAGTGTACGACTTCCAGTAATCTTCAAGGATCAGCC 1216
 Db 321 LysAsnProValValTyrAlaValPheThrThrSerSerAsnIlePheLysGlySerAla 340
 QY 1217 GTGTGATATGATGATGATGATGTGAGAGGGTGTCTTGGTCCATATGCCACAGG 1276
 Db 341 ValCysMetTyrSerLeuThrAspIleArgArgValPheLeuGlyProTyrAlaHisArg 360
 QY 1277 GATGACCCCACTATCAATGGTGCCTTATCAAGNAGAGTCCCTATCCACGGCCAGA 1336
 Db 361 AspGlyProAsnTyrGlnTrpValProPheGlnGlyArgValProTyrProArgProGly 380
 QY 1337 ACTTGTCCTCCAGCAAAACATTTGGTGGTTTGACTCTACAAAGGACCTTCTCTGATGTT 1396
 Db 381 ThrCysProArgGlnThrPheGlyGlyPheAspSerThrLysAspLeuProAspGluVal 400
 QY 1397 ATAACTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTGTCTTATGAACATCGC 1456
 Db 401 IleMetPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArg 420
 QY 1457 CCAATAGTGTCAAAAACGGATGTAATATCAATTTACACAAATTCGTAGACCGAGTG 1516
 Db 421 ProIleIleIleLysThrGluValAspTyrGlnPheThrGlnIleValValAspArgVal 440
 QY 1517 GATCAGAGAAGTGCACAGTATGATGTTATCGAAACAGATGTTGGGACCCCTTCTT 1576
 Db 441 GluAlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspMetGlyThrIleLeu 460
 QY 1577 AAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCTGGAAGAA 1636
 Db 461 LysValValSerValProLysGluThrTrpThrAspLeuGluGluValLeuLeuGluGlu 480
 QY 1637 ATGACAGTTTTCGGGACCGACTGCTATTTTCACAATGGAGCTTCCACATGAAGCAGCAA 1696
 Db 481 MetThrValPheArgGluProThrAlaIleSerAlaMetGluIleSerThrLysGlnGln 500
 QY 1697 CAACTATATATTGTTCAACCGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATT 1756
 Db 501 GlnLeuTyrIleGlySerSerValGlyValSerGlnLeuProLeuHisArgCysAspVal 520
 QY 1757 TACGGGAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTTGGATCGT 1816
 Db 521 TyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 540
 QY 1817 TCTGCATGTTCTCGCTATTTCCTCCACTGCAAGAGACGCAACAGACAGATATAGA 1876
 Db 541 SerSerCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArg 560
 QY 1877 AATCGAGACCCACTGACTCTGTTTCAGACTTA---CACCATGATATATCACCATGGCCAC 1933
 Db 561 AsnGlyAspProLeuThrHisCysSerAspLeuGlnHisGlnAspAspProHisArgGln 580
 QY 1934 AGCCCTGAAGAGAGAATCATCTATGGTGTAGAGAAATAGTAGACATTTTGGATGCACT 1993
 Db 581 SerLeuGluGluLysIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSer 600
 QY 1994 CCGAAGTCGACAGAGCGCTGCTATTGSCAAATTCAGAGCGCAAAATGAGAGCGAAA 2053
 Db 601 ProLysSerGlnArgAlaLeuValPheTrpGlnPheGlnLysGlnAsnGluGluLysLys 620
 QY 2054 GAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCTTCTCTAGTACT 2113
 Db 2113

```

Db      621 AspGluLeuValAspGluArgLysIleLysThrGluHisGlyLeuLeuLeuArgThr 640
QY      2114 CTACACAGAGGATTTCAGCAATACCTCTCCATGCGGTGGAGCATGGGTTCATACAA 2173
Db      641 LeuLysLysArgAspSerGlyIleTyrTyrCysAsnAlaValGluHisGlyPheMetGln 660
QY      2174 ACTCTCTTAAGGTAACTGACCTGAAGTCATTCACAGAGCATTTGGAAGAACTCTTCAT 2233
Db      661 ThrLeuLeuLysValThrLeuGluIleLysThrGluHisGlyLeuAspGluLeuLeuHis 680
QY      2234 AAAGATGATGAT---GGAGATGGCTCTAAGACCAAGAAATCTCAATAGCATGACACT 2290
Db      681 LysGluAspGluGlyAspSerHisLysHisLysGluProSerAsnSerMetSerPro 700
QY      2291 ACCGAGAAGCTGTGTACAGACATTCATGACAGCTCATCAACACCCCAATCTCAACACG 2350
Db      701 ThrGlnLysIleTyrTyrArgAspPheMetGlnLeuLeuAsnHisProAsnLeuAsnThr 720
QY      2351 ATGGATGAGTCTGTGACACAGTTTGGAAAGGGACCGGAAACAACAGCTCGGCAAGGCCA 2410
Db      721 MetAspGluPheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnLysAsn 740
QY      2411 GGACATACCCCGAGGAAACAGTAACAAATGGAGACACTTACAGAGAAATAAGAAAGGTAGA 2470
Db      741 GlyAsnValGlnValSerAsnThrLysIleLysHisLeuGlnGluAsnLysLysGlyArg 760
QY      2471 AACAGGAGACCCACCAATTTGAGAGGCGCACCCAGGAGTGTC 2512
Db      761 AsnArgArgThrHisGluPheGluArgAlaProArgSerVal 774

RESULT 2
QBH34
ID      QBH34      PRELIMINARY;      PRT;      777 AA.
AC      QBH34;
DT      01-MAR-2003 (TREMREL. 23, Created)
DT      01-MAR-2003 (TREMREL. 23, Last sequence update)
DT      01-MAR-2003 (TREMREL. 23, Last annotation update)
DE      Semaphorin 3D precursor homolog.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=1090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Kidney, and Skin;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium.
RA      The RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RL      60,770 full-length cDNAs."
RL      Nature 420:563-573 (2002).
DR      EMBL; AK028900; BAC36185.1; -.
DR      EMBL; AK052671; BAC35092.1; -.
SQ      SEQUENCE 777 AA; 89548 MW; 5450D8D45D1BDABF CRC64;

Alignment Scores:
Pred. No.:      4,346-183      Length:      777
Score:      2258.00      Matches:      428
Percent Similarity:      72.24%      Conservative:      147
Best Local Similarity:      53.77%      Mismatches:      190
Query Match:      46.12%      Indels:      31
Db:      11      Gaps:      13

US-09-774-490-1 (1-2709) x QBH34 (1-777)
QY      148 ACNATACAGGAGGAGACTAAGACAGCAAGGGACCTACAGCGTCTCCAGCATGGGCTG 207
Db      4 ThrLysAspGluAsnProArgSerArgSerGlnAspLeu---HisLeuPheHis-AlaTr 22
QY      208 GTTAACTAGGATGTC---TGCTTTCTGGGAGTATTACTTACAGCAAGGCAAACTA 264
Db      22 pMetLeuLeuMetThrValLeuPheLeuProValThrGluThrSer----- 38

```

```

QY      265 TCAGAAATGGGAAGAACAAATGTGCCAAGCTGAAATTTATCTCACAAGAAATGTTGGAATC 324
Db      39 -----LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSe 55
QY      325 CAACAATGTGATCATTCTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTCTCTTTT 384
Db      55 rAsnThrCysIleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuSe 75
QY      385 GGATGAGGAACCGAGTAGCTGTATGTGGAGCAAGAGTACATATTTTTCATTTTCATTCGACT 444
Db      75 uAspGluGluArgGlyLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuLeuSe 95
QY      445 GGTTAATATC---AAGGATTTTCAAAAGATTGTGTGCCAGTATCTTACACAGAGAGA 501
Db      95 uValAspLeuAsnLysAsnPheLysIleTyrTyrProAlaAlaLysGluArgValG1 115
QY      502 TGAATGCAAGTGGCTGGAAGACATCTCTGAAGAATGTCTATTTTCATCAAGTACT 561
Db      115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgVal 135
QY      562 TAAGGCATATATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTTCATCAATTTG 621
Db      135 uGlnProTyrAsnLysThrHisValTyrValCysGlyThrGlyAlaPheHisProLeuCy 155
QY      622 CACCTACATTTGAATTTGGACATCATCTCGAGAGCAATATTTTAAGCTGGAGAACTACA 681
Db      155 sGlyTyrIleAspLeuGlyAlaAsnLysGluLysLysLeuAspThrHisAs 175
QY      682 TTTTGAACCGCGTGGAGAGTCCATATGACCTAAGCTGTGTGACAGCATCTCTTTT 741
Db      175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerVal 195
QY      742 AATAGATGGAGAAATATATCTCTGGAACCTGACGTGATTTTATGGGGCGAGACTTTGCTAT 801
Db      195 tThrAspGluHisLeuTyrSerGlyThrAlaSerAspPheLeuLysLysAspThrAlaPh 215
QY      802 CTTCGAACTCTCGG-----CACCAACCCCAATCAGACAGACAGCAGCATGATTC 852
Db      215 eThrArgSerLeuGlyLeuMetGlnAspHisHisSerIleArgThrAspIleSerGluHi 235
QY      853 CAGTGGCTCAATGATCCAAAGTTTCATAGTGGCCACCTCATCTCAGAGAGTGACAATCC 912
Db      235 sHisTrpLeuAsnGlyAlaLysPheIleGlyThrPheProIleProAspThrTyrAsnPr 255
QY      913 TGAAGATGACAAAGTATATCTTTCTTCGCGGAAATGCAATAGATGAGAGCACTCTGG 972
Db      255 oAspAspLysIleTyrPhePheArgLysSerSerGlnGluGlySerThrSerAs 275
QY      973 AAAAGCTACTCACGCTAGAATAGCTCAGATATGCAAGAATGACTTTTGAGGGGCACAGAAG 1032
Db      275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSe 295
QY      1033 TCTGTGAATTAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTC 1092
Db      295 rLeuIleAsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY      1093 AAATGGCATTCACACTCATTTTGTAGTAACCTGAGATGTATTCTCTAATGAATTTAAAGA 1152
Db      315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAs 335
QY      1153 TCCTAAATAATCCAGTTGTATAGGAGTGTTCACGATTCACAGTAACTTTTCAAGGGATC 1212
Db      335 pGluArgAsnProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySe 355
QY      1213 AGCGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATGCCCCA 1272
Db      355 rAlaValCysValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHi 375
QY      1273 CAGGATGGAGCCCAACTATCAATGGGGCTTATCAAGGAAGAGTCCCTTATCCACGCCC 1332
Db      375 sLysGluSerAlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgPr 395
QY      1333 AGGAACCTTGTCACAGCAAAACATTTTGTGGT---TTTGACTCTACAAAGGACCTTCCTGA 1389

```

Db 395 ogyThrCysProSerLysThrTyAspProLeuLeuLysSerThrArgAspPheProAs 415
QY 1390 TGATGTATAACCTTGGCAAGATCATCCAGCCATGTACATCCAGTGTTCCTATGAA 1449
Db 415 pAspValIleSerPheIleArgHisProValMetTyrLysSerValTyProValAl 435
QY 1450 CAATCGCCCAATAGTATCAAAACCGATGTAATATCAATTTACAAATTTGCTAGA 1509
Db 435 aglyAlaProThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValAs 455
QY 1510 CCCAGTGTATGAGAGATGAGCAGTATGATGTATGTTATTCGGAACAGATGTTGGAC 1569
Db 455 phisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyTh 475
QY 1570 CTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGTTCCT 1629
Db 475 rValLeuLysValSerIleSerLysGlyLysTrp--AenMetGluGluValValle 494
QY 1630 GGAAGAAATGACAGTTTTCGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAA 1689
Db 494 uGluGluLeuGlnValPheLysHisProThrAlaIleLeuAenMetGluLeuSerLeuLy 514
QY 1690 GCAGCAACAACTATATTTGTTTCAACGGCTGGGTGCCAGCTCCCTTTACACGGTG 1749
Db 514 sgGlnGlnLeuTyrValGlySerTrpAspGlyLeuValGlnLeuSerLeuHisArgCy 534
QY 1750 TGATATTACGGAAGCGTGTGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 1809
Db 534 sAspThrTyrGlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTr 554
QY 1810 GGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869
Db 554 pAspGlyAenAlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAs 574
QY 1870 TATAAGAAATGGAGACCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929
Db 574 pValLysTyrGlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis-- 593
QY 1930 CCACAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAGACATATTTTGAATG 1989
Db 594 -GluThrAlaAspGluLysValIlePheGlyIleGluPheAenSerThrPheLeuGluCy 613
QY 1990 CAGTCCGAGTCCGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2049
Db 613 sileProLysSerGlnGlnAlaSerValGluTrpTyrIleGlnArgSerGlyAspGluHi 633
QY 2050 AAAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2109
Db 633 sArgGluGluLeuLysProAspGluArgIleIleLysThrAspTyrGlyLeuLeuIleAr 653
QY 2110 TAGTCTACAAAGAGATTCAGGATTAATCTCTGCGCATGCGGTGGGAACATGGTTCAT 2169
Db 653 gSerLeuGlnLysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheI 673
QY 2170 ACAAACTCTTCTAGGTAACTCGAAGTCAATGACAGAGCATTTGGAGAACTTCT 2229
Db 673 ehistrIleValLysLeuThrLeuAsnValIleGluAenGluGlnMetGluAsn--Th 692
QY 2230 TCATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2289
Db 692 rGlnArgAlaGluTyrGlnGluGlyGlnValLysAspLeuLeuAlaGluSer----- 709
QY 2290 TAGCCAGAAAGTCTGGTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACAC 2349
Db 710 -----ArgLeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe--Se 726
QY 2350 GATGATGATGCTGTGGAACAAGTTTGGAAAGGAGCGAAACACGTCGGCAAGGCC 2409
Db 726 rIleuAspGlnTyrCysGluGlnMetTyrTyrLysGluLysArgGlnArgAsnLys-- 745
QY 2410 AGGACATACCCCGGGAACAGTAACAAATGGAAGCATTACAAAGAAATAAGAAAGGTAG 2469

746 -----GlySerProLysTrpLysHisMetGlnGluMetLysLysLysAr 760
QY 2470 AAACAGGAGGACCCAC---GAATTTGAGAGCGCCACCCAGGAGTCTC 2512
Db 760 gasnArgArgHisArgAspLeuAspGluLeuGlnArgSerVal 775
RESULT 3
Q8BMP6
ID Q8BMP6 PRELIMINARY; PRT; 777 AA.
AC Q8BMP6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Senaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK031704; BAC27522.1; --
SQ SEQUENCE 777 AA; 89562 MW; B7204D82288B89CE CRC64;
Alignment Scores:
Pred. No.: 1,48-182 Length: 777
Score: 2252.00 Matches: 427
Percent Similarity: 72.11% Conservative: 147
Best Local Similarity: 53.64% Mismatches: 191
Query Match: 46.00% Indels: 31
Gaps: 13
US-09-774-490-1 (1-2709) x Q8BMP6 (1-777)
QY 148 ACAATACAGGAAGAGCAATTAAGCAAGGAGCAAGGAGCTGTGACAGCTGGGCTG 207
Db 4 ThrLysAspGluAsnProArgSerArgSerGlnAspLeu---HisLeuPheHis-AlaTr 22
QY 208 GTTAACAGATTGTC---TGTCTTTCTGGGAGTATTACTTACAGCAAGCAAACTA 264
Db 22 pMetMetLeuIleMetThrValLeuPheLeuProValThrGluThrSer----- 38
QY 265 TCAGAAATGGGAAGCAATGTGCCAAGCTGAAATTTATCTTACAAAGAAATGTTGGAATC 324
Db 39 -----LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSe 55
QY 325 CAACAATGTGATCACTTCAATGCTTGGCAAGCTCCAGTTCAGTTATCATCTCTCTTT 384
Db 55 rAsnThrCysIleProPheLeuGlySerSerGlyLeuAspPheGlnThrLeuLeuLe 75
QY 385 GGATGAGAAACGAGTAGGCTGTATGTTGGCAAGAGATCATATTTTTCATTCGACT 444
Db 75 uAspGluGluArgGlyLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuSerLe 95
QY 445 GTTTAAATATC---AAGGATTTTCAAAAGATTTGTGCGCAGTATCTTACACCAAGAGA 501
Db 95 uValAspLeuAsnLysAsnPheLysLysIleTyrTrpProAlaLysGluArgValG 115
QY 502 TGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGATGTCTTAATTTTCAATCAAGGTACT 561
Db 115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgVal 135
QY 562 TAAGGCATATATCAGACTCATCTTGTACGCTGTGGACCGGGCTTTTCAATCAATTTG 621
Db 135 uGlnProTyrAsnLysThrHisValTyrValCysGlyThrGlyAlaPheHisPLeuLeu 155
QY 622 CACCTACATTGAAATTTGACATCATCTCTGAGGACAATATTTTAAAGCTGGAGAACTACA 681

Db 155 sGlyTyrIleAspLeuGlyAlaAsnLysGluGlnLeuIlePheLysLeuAspThrHisAs 175
QY 682 TTTTGAAGCGCGTGGAGAGTCCATATGACCTTACAGCTGCTGACAGACCTCCCTTTT 741
Db 175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMe 195
QY 742 AATAGATGAGAAATATATCTCTGGAACCTGACCTGATTTATGCGGCGAGACTTGGCTAT 801
Db 195 tThrAspGluHisLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPh 215
QY 802 CTTCCGAACCTCTGGG-----CACCACCAACCAATCAGGACAGACGACGATGTC 852
Db 215 eThrArgSerLeuGlyLeuMetGlnAspHisHisSerIleArgThrAspIleSerGluHi 235
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCACCTCATCTCAGAGAGTCAACATCC 912
Db 235 shiStrLeuAenGlyAlaLysPheIleGlyThrPheProIleProAspThrTyrAsnPr 255
QY 913 TGAAGATGACAAAGTATATCTTTCTCGTGAAGAAATGCAATGATGAGAACACTCTGG 972
Db 255 oAspAspAspLysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAs 275
QY 973 AAAAGCTACTCAGCTAGATAGTGCAGATATGCAAGATGACTTGGAGGCGACAGAG 1032
Db 275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSe 295
QY 1033 TCTGGTGAATAATGCAACATCTCTCAAAGCTCGTGTGATTTCTCAGTCCAGCTGCC 1092
Db 295 rIeuIleAenLysTyrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY 1093 AAATGGCATGTACATCTATTTGATGAATCGAGGATGATTCCTTAATGAATTTAAAGA 1152
Db 315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAs 335
QY 1153 TCCTAAAATCCAGTTGATATGAGTGTGTACAGCTTCCAGTATCCATGATTCACAGGATC 1212
Db 335 pGluArgLysProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySe 355
QY 1213 AGCCGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272
Db 355 rAlaValCysValTyrSerMetAlaAspIleArgAlaValPheAenGlyProTyrAlaHi 375
QY 1273 CAGGATGAGACCAACTATCAATGAGTGCCTTATCAAGAGAGAGTCCCTATCCAGGCC 1332
Db 375 sLysGluSerAlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgPr 395
QY 1333 AGGAATCTTCCAGCAAAACATTTGGTGGT---TTTGACTCTACAAAGGACCTTCCTGA 1389
Db 395 oGlyThrCysProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAs 415
QY 1390 TGATGTTTAACTTTGCAAGAGTATCCAGCCATGATACATCCAGTGTTCCTATGAA 1449
Db 415 pAspValIleSerPheIleArgArgHisProValMetTyrLysSerValTyrProValAl 435
QY 1450 CAATCGCCCAATGATGATCAAAACCGAGTGTAAATATCAATTTACAAATTTGCTGAGA 1509
Db 435 aclyAlaProThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAs 455
QY 1510 CCGAGTGAATGAGAGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1569
Db 455 pHisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyTh 475
QY 1570 CQTCTCTAAAGTAGTTTCAATTTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCT 1629
Db 475 rValLeuLysValValSerIleSerLysGluLysTyr---AsnMetGluGluValValLe 494
QY 1630 GGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTACAGCAATGGAGCTTTTCCACTAA 1689
Db 494 uGluGluLeuGlnValPheLysHisProThrAlaIleLeuAsnMetGluLeuSerLeuLy 514
QY 1690 GCAGCAACACTATATATTTGTTCAACCGCTGGGTGCTCCAGCTCCCTTTACACCGGTG 1749

Db 514 sGlnGlnGlnLeuTyrValGlySerTyrAspGlyLeuValGlnLeuSerLeuHisArgCy 534
QY 1750 TGATATTTACCGGAAAGCGTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1809
Db 534 sAspThrTyrGlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTr 554
QY 1810 GGATGGTCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCACAGACGACAAGA 1869
Db 554 pAspGlyAsnAlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAs 574
QY 1870 TATAAGAAATCGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATACCATGG 1929
Db 574 pValLysTyrGlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis-- 593
QY 1930 CCACAGCCCTCAAGAGAGAATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATG 1989
Db 594 -GluThrAlaAspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCy 613
QY 1990 CAGTCCGAGTCCGAGAGAGCGCTGCTTATGGAATTCAGAGCGCGAAATGAAGAGG 2049
Db 613 sIleProLysSerGlnGlnAlaSerValGluTyrTyrIleGlnArgSerGlyAspGluHi 633
QY 2050 AAAAGAGAGATCAGATGAGTATCATCATCAGGACAGATCAAGGCTTCTGCTAGC 2109
Db 633 sArgGluGluLeuLysProAspGluArgIleIleLysThrAspTyrGlyLeuLeuIleAr 653
QY 2110 TAGTCTCAACAGAGATTCAGGCAATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169
Db 653 sSerLeuGlnLysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIl 673
QY 2170 ACAACTCTTCTTAAGTAACCTGGAAGTCAATTCAGACAGAGCATTTGGAGAACTTCT 2229
Db 673 eHisThrIleValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsn---Th 692
QY 2230 TCATAAGATCATGATGAGATGCTCTAAGACCAAGAAATGTCCATGATGATGATGATG 2289
Db 692 rGlnArgAlaGluTyrGlnGluGlyGlnValLysAspLeuLeuAlaGluSer----- 709
QY 2290 TAGCCAGAGCTGCTGACAGAGACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2349
Db 710 -----ArgLeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---Se 726
QY 2350 GATGATGAGTCTGTTGAACAAAGTTTGAAGAGGAGCGGAAACCACTCGGCAAGGCC 2409
Db 726 rLeuAspGlnTyrCysGluGlnMetTyrTyrLysLysLysArgGlnA-GasnLys-- 745
QY 2410 AGGACATACCCCGGAGACACTAACAAATGCAAGCTTACAGAAATAAGAAAGGTAG 2469
Db 746 -----GlySerProLysTyrLysHisMetGlnGluMetLysLysLysLeAr 760
QY 2470 AACAGAGGAGGCCAC---GAATTTGAGAGGCGGCCACCCAGGAGTGTCT 2512
Db 760 gasnArgArgHisHisArgAspLeuAspGluLeuGlnArgSerVal 775

RESULT 4
Q8TB71 PRELIMINARY; PRT; 748 AA.
AC Q8TB71
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024220; AAB24220.1; -;
InterPro; IPR003599; Ig.

DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM01403; Sema; 1.
 DR SMART: SM00409; Ig; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS00835; IG_LIKE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 748 AA; 83034 MW; 587C53CB65AB4656 CRC64;

Alignment Scores:
 Pred. No.: 3, 29e-170 Length: 748
 Score: 2106.50 Matches: 404
 Percent Similarity: 67.09% Conservative: 122
 Best Local Similarity: 51.53% Mismatches: 185
 Query Match: 43.02% Indels: 73
 DB: 4 Gaps: 13

US-09-774-490-1 (1-2709) x Q8TB71 (1-748)

QY 221 GTCGTCTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTATCAGAATGGGAAGAC 280
 DB 13 LeuAlaLeuLeuTyrAlaValGlyLeuGlySerAlaAla-----Pro 26
 QY 281 ANTGGCCAGGCTGAATATTCCTACAAAGAAATGTTGGAATCCAACTGATCACT 340
 DB 27 SerProArgLeuArgLeuSerPheGlnGluLeuGlnAlaTrpHisGlyLeuGlnThr 46
 QY 341 TTCATGGCTCGCCCAACAGCTCCAGTTATCATCTTCTTTCGATGAGCAAGCGAGT 400
 DB 47 PheSer---LeuLuarGthrCysCysTyrGlnAlaLeuLeuValAspGluGly 65
 QY 401 AGGCTGTATGTTGGCAAGAGTACATATTTTCATTCGACCTGGTTAATATC---AAG 457
 DB 66 ArgLeuPheValGlyAlaGluAenHisValAlaSerLeuAenLeuAsnLeuSerLys 85
 QY 458 GATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGT 517
 DB 86 ArgAlaLysLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
 QY 518 GGAAGAGACATCTGAAGAAGTGTCTAATTTTCATCAAGTACTTAAGSCATATATCAG 577
 DB 106 GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaTyrAsnArg 125
 QY 578 ACTCACTGTAGCCCTGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATGAAAT 637
 DB 126 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal 145
 QY 638 GGCATCATCTGAGACAAATATTTTAAAGCTGGAGAACTACATTTTGAACCGCGT 697
 DB 146 GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys 165
 QY 698 GGGAGAGTCCATATGACCTAAGCTGCTGACAGCATCTCTTTAATAGATGAGAATTA 757
 DB 166 GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValGlyGluGluLeu 185
 QY 758 TACTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTGCTATCTTCCGAACCTCTGGG 817
 DB 186 TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly 205
 QY 818 CACCACCCCAATCAGGACAGAGCATGATTCAGGTGGCTCATGATCAAAAGTTC 877
 DB 206 GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGluProLysPhe 225
 QY 878 ATTAGTGGCCACTCATCTCAGAGTGTACATCTCTGAAGATGACAAAGTATCTTTTC 937
 DB 226 ValLysValPheTrpIleProGluSerGluAenProAspAspLysIleTyrPhePhe 245
 QY 938 TTCCTGGAATATGCAATAGATGGAGAACAACCTCT---GGAAAGCTACTCACCTAGAATA 994
 DB 938 LeuGlnAlaArgValGluTrpThrPheGlnArgAlaGlyValThrAlaHisThrGlnVal 615

DB 246 PheArgGluThrAlaValGluAlaProAlaLeuGlyArgLeuSerValSerArgVal 265
 QY 995 GGTGAGATATCAAGAAATGACTTTGGAGGGCACAGAGTCTGGTCAATAAATGGAACA 1054
 DB 266 GlyGlnIleCysArgAsnAspValGlyGlyGlnArgSerLeuValAsnLysTrpThr 285
 QY 1055 TTCCTCAAAGCTCGTCTGATTGCTCAGTGCAGGTCCAAATGGCATTGACACTCATTTT 1114
 DB 286 PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe 304
 QY 1115 GATCAATGTCAGGATGATTCTTAATGAACCTTTAAAGATCTCAAAATCCAGTGTATAT 1174
 DB 305 AspGlnLeuGlnAspValPheLeuLeuSerSerArgAspHisArgThrProLeuLeuTyr 324
 QY 1175 GGAGTGTTCAGACTTCAGTAACATTTTCAAGGATCAGCCGTGTGATGATPATGACATG 1234
 DB 325 AlaValPheSerThrSerSer---IlePheGlnGlySerAlaValCysValTyrSerMet 343
 QY 1235 AGTGAATGAGAAGGGTGTCTTCTTGGTCCATATGCCACAGGGATGGACCAACTATCAA 1294
 DB 344 AsnAspValArgAlaPheLeuGlyProPheAlaHisLysGluGlyProMetHisGln 363
 QY 1295 TGGTGCCTTATCAAGAAAGAGTCCCTATCCAGGCCAGGAACCTTGTCCAGCAAAACA 1354
 DB 364 TrpValSerTyrGlnGlyArgValProTyrProArgProGlyMetCysProSerLysThr 383
 QY 1355 TTTGGTGGTGTGACTCTACAAAGGACCTTCTCATGATGTTATATACTTTGCAAGAAGT 1414
 DB 384 PheGlyThrPheSerSerThrLysAspPheProAspAspValIleGlnPheAlaArgAsn 403
 QY 1415 CATCAGGCATGTACAATGCTTCTTATGAACATCCCAATCGCCCAATAGTAGTCAAAACG 1474
 DB 404 HisProLeuMetTyrAsnSerValLeuProThrGlyGlyArgProLeuPheLeuGlnVal 423
 QY 1475 GATGTAATATCAATTTTACAAATGCTCTAGACGAGTGGATGCGAGAGATGAGACAG 1534
 DB 424 GlyAlaAsnTyrThrPheThrGlnIleAlaAlaAspArgValAlaAlaAlaAspGlyHis 443
 QY 1535 TATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAAAGTAGTTTCAATTCCT 1594
 DB 444 TyrAspValLeuPheIleGlyThrAspValGlyThrValLeuLysValIleSerValPro 463
 QY 1595 AAGGAGACTGCTGATGATTAGAAAGAGTCTCTGCTGGAAGAAATGACAGTTCCTCGGAA 1654
 DB 464 LysGlySerArgProSerAlaGluGlyLeuLeuGluLeuHisValPheGluAsp 483
 QY 1655 CCGACTGCTATTTCAGCAATGAGCTTTCCTACTAAGCAGCAACAACTATATTTGCTTCA 1714
 DB 484 SerAlaAlaValThrSerMetGlnIleSerSerLysArgHisGlnLeuTyrValAlaSer 503
 QY 1715 ACGCTGGGTTCGCCAGCTCCCTTACACCGGTGTGATATTTACGGGAAAGCGTGTGCT 1774
 DB 504 ArgSerAlaValAlaGlnIleAlaLeuHisArgCysAlaAlaHisGlyArgValCysThr 523
 QY 1775 GAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGGATGGTTCGTGCTATGTTCTCGCTAT 1834
 DB 524 GluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyValAlaCysThrArgPhe 543
 QY 1835 TTTCCCTCAAGAGAGACGACAGCAGCAGCAAGATATAGAAATGAGAGCCACTGACT 1894
 DB 544 GlnProSerAlaLysArgArgPheArgGlnAspValArgAsnGlyAspProSerThr 563
 QY 1895 CACTGTTTCAGACTTACACCATGATAATCACCATGCCACAGCCCTGAA----- 1942
 DB 564 LeuCysSer-----GlyAspSerSerArgProAlaLeuLeu 575
 QY 1943 GAGAAATCATCTATGTTGTAGAAATAGTAGCACATTTTGGATGCGAGTCCGAAGTCG 2002
 DB 576 GluHisLysValPheGlyValGluGlySerSerAlaPheLeuGluCysGluProArgSer 595
 QY 2003 CAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGGAATAGAGCGGAAAGAGAGATC 2062
 DB 596 LeuGlnAlaArgValGluTrpThrPheGlnArgAlaGlyValThrAlaHisThrGlnVal 615

```
QY 2063 AGAGTGGATGATCATATCATCATGAGCAGATCAAGGCTTCTCTAGCTAGTCTACAACAG 2122
Db 616 LeuAlaGluGluArgThrGluArgThrAlaArgGlyLeuLeuLeuArgLeuArgArg 635
QY 2123 AAGGATTCAGGCAATACCTCTGCGATCGGTTGGACATCGGTTTCATCAAACTCTTCT 2182
Db 636 ArgAspSerGlyValTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 655
QY 2183 AAGGTAACCTCGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTTCATAAAGATGAT 2242
Db 656 ArgLeuSerLeuHisValLeuSerAlaThrGlnAlaGluArgLeu----- 670
QY 2243 GATGAGATGCTCTAAGACCAAGAAATGTCATATAGCATGACACCTAGCCAGAGGTC 2302
Db 671 -----AlaArgAlaGluGluAlaAlaProAlaAlaProProGlyProGlyLeu 686
QY 2303 TGGTACAGAGCTTCATGCGAGCTCATCAACACCCCACTTCAACACAGTGGATGATC 2362
Db 687 TrpTyrArgAspPheLeuGlnLeuVal----- 695
QY 2363 TGTGAACAAGTTTGGAAAGGACCGGAAACAAACGTCGGCAAGGCCAGCATACCCCA 2422
Db 696 -----GluProGly-----GlyGly 700
QY 2423 GGGACAGTAAACAAATGGAAG-----CACTTA 2449
Db 701 GlySerAlaAsnSerLeuArgMetCysArgProGlnProAlaLeuGlnSerLeuProLeu 720
QY 2450 CAAGAAATAGAAAGGTAGAACAGGAGGAGGCCAC-----GAATTGAGAGG 2497
Db 721 GluSerArgArgGlyGlyArgAsnArgArgThrHisAlaProGluProArgAlaGluArg 740
QY 2498 GCACCCAGGAGT 2509
Db 741 GlyProArgSer 744

RESULT 5
Q8TDV7
ID Q8TDV7 PRELIMINARY; PRT; 754 AA.
AC Q8TDV7;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Semaphorin 3B.
GN SEMA 3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Koyama N.;
RT "semaphorin 3B (SEMA3B) CDNA.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083186; BAB88870.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; SEMA.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF01403; SEMA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; SEMA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 754 AA; 83691 MW; E61FD08C04E9A68E CRC64;

Alignment Scores:
Pred. No.: 7, 21e-170 Length: 754
Score: 2102.50 Matches: 403
Percent Similarity: 66.79% Conservative: 124
```

```
Best Local Similarity: 51.08% Mismatches: 185
Query Match: 42.94% Indels: 77
DB: 4 Gaps: 13

US-09-774-490-1 (1-2709) x Q8TDV7 (1-754)

QY 221 GTCTGCTCTTTCTGGGAGATTAATTACAGCAGAGCAAACTATCAGAAATGGGAGAAC 280
Db 13 LeuAlaLeuLeuTrpAlaValGlyLeuGlySerAlaAla-----Pro 26
QY 281 AATGTGCCAAGCTGAAATATCTTACAAAGAAATGTTGGAATCCAACTGATGATCACT 340
Db 27 SerProProArgLeuArgLeuSerPheGlnGluLeuGlnAlaTrpHisGlyLeuGlnThr 46
QY 341 TTCAATGCTTGGCCCAACAGCTCCAGTTTATCATACCTTCTTTTGGATGAGGACGGAGT 400
Db 47 PheSer---LeuGluArgThrCysCysTyrGlnAlaLeuValAspGluGluArgGly 65
QY 401 AGGCTGTATGTTGGAGCAAGGATCACAATTTTTCATTCGACCTGGTTAATATC---AAG 457
Db 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuAsnLeuAspAsnIleSerLys 85
QY 458 GATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAGAGAGATGAATGCAAGTGGGCT 517
Db 86 ArgAlaLysLysLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
QY 518 GGAAAGACATCTCGAAAGAAATGTCTAATTTTCATCAAGGTACTTAAAGCATATATACAG 577
Db 106 GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaTyrAsnArg 125
QY 578 ACTCATCTGTACGCTGTGGAAACGGGGCTTTTCATCAATTTGACCTACATGAAAT 637
Db 126 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal 145
QY 638 GGACATCATCTCGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACACGGCGT 697
Db 146 GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys 165
QY 698 GGGAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATTA 757
Db 166 GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValGlyGluLeu 185
QY 758 TACTCTGGAATCGAGCTGATTTTATGGGCGAGACTTTTGTATCTTCGAACTCTTGG 817
Db 186 TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly 205
QY 818 CACCACCAACCAATCAGCAGACGAGCATGATTCAGCTGGCTCAATCATCCAAAGTTC 877
Db 206 GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGluProLysPhe 225
QY 878 ATTAGTGCCCACTCATCTCAGAGAGTGCACAACTCTGAAGATGACAAAGTATATCTTTTC 937
Db 226 ValLysValPheTrpIleProGluSerGluAsnProAspAspAspValIleTyrPhePhe 245
QY 938 TTCCGTGAAATGCAATAGATGGAGAACACTCT---GGAAAGCTTACTCACGCTGAATA 994
Db 246 PheArgGluThrAlaValGluAlaAlaProAlaLeuGlyArgLeuSerValSerArgVal 265
QY 995 GGTGAGATATCAAGAATGATTTGGAGGGCAGACAGACTCTGGTCAATTAATGACACACA 1054
Db 266 GlyGlnIleCysArgAsnAspValGlyGlnArgSerLeuValAsnLysTrpThr 285
QY 1055 TTCTCTCAAGCTCGTGTATTTGTCTAGTGCAGCTCCAGCTCCAAATGGCATTGACATCTTTT 1114
Db 286 PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe 304
QY 1115 GATGAACCTG-----CAGGATGTATTCTTAATGAATTTTAAAGATCTCTAAA 1159
Db 305 AspGlnLeuArgProPheProAlaGluAspValPheLeuLeuSerSerArgAspHisArg 324
QY 1160 AATCCAGTTGTATATGAGGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db 325 ThrProLeuLeuTyrAlaValPheSerThrSerSerSerThrGlnGlySerAlaVal 344
```


Db 759 LeuArgSerLysAlaGluHisPheArgLeuProArg 770

RESULT 7

Q8QUG9 PRELIMINARY; PRT; 756 AA.

AC Q8QUG9; 01-JUN-2002 (trEMBLrel. 21, Created)

DT 01-JUN-2002 (trEMBLrel. 21, Last sequence update)

DT 01-MAR-2003 (trEMBLrel. 23, Last annotation update)

DE Semaphorin 3F.

GN SEMA3F.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OC NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RT Watanabe Y., Nakamura H.;

RT "Axon guidance of the trochlear nerve by Sema 3F along mid-hindbrain boundary.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB072930; BAB8691.1; -.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG_c2.

DR InterPro; IPR003006; IG_MHC.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR001627; Sema.

DR Pfam; PF00047; ig; 1.

DR Pfam; PF01403; Sema; 1.

DR SMART; SM00408; IGc2; 1.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; Sema; 1.

DR PROSITE; PS08335; IG_LIKE; 1.

KW Immunoglobulin domain.

SQ SEQUENCE 756 AA; 85196 MW; C36754C02541ED88 CRC64;

Alignment Scores:

Pred. No.: 128e-161 Length: 756

Score: 2005.50 Matches: 380

Percent Similarity: 68.16% Conservative: 138

Best Local Similarity: 50.00% Mismatches: 217

Query Match: 40.96% Indels: 25

DB: 13 Gaps: 11

US-09-774-490-1 (1-2709) x Q8QUG9 (1-756)

QY 227 CTTTCTGGGAGTATTACTACAGCAGAGCAAACTATCAGAATGGGAAGAACAAATGTG 286

Db 7 LeuLeuTriaAlaThrLeuLeuThrLeuGlyTyrArgAlaAlaHisGlyLysAspGlyVal 26

QY 287 -----CAGGCTGAATATCTCTACAAAGAAATGTTGGAATCCAAATGTGATC 337

Db 27 ProProThrProArgValGlnLeuSerPheLysGluLeuLysAlaThrGlyThrAlaHis 46

QY 338 ACTTCAAGGCTGGCCACAGCTCCAGTTATCATACCTTCTTTGGATCAGGAACGG 397

Db 47 PhePheAsnPheLeuLeuAsnSerAspTyrArgIleLeuLeuLysAspGluAspHis 66

QY 398 AGTAGCTGTATGTTGGCAAGAGATCACATATTTTCATTCGACCTGTTAATATCAAG 457

Db 67 AspArgMetTyrValGlySerLysAspTyrValLeuSerLeuAspLeuHisAspIleAsn 86

QY 458 GATTTTCAAAAGATTGTG---TGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514

Db 87 ArgGluProLeuIleIleHisTyrProAlaSerGlnGlnArgIleGluGluCysIleLeu 106

QY 515 GCTGGAAGAGATCCTGAAAGAAATGCTGAATTTTCATCAAGTACTTAAGGCATATAAT 574

Db 107 SerGlyLysAsnSerAsnGlyGluCysGlyAsnPheIleArgLeuIleGlnProtrpAsn 126

QY 575 CAGACTCATTGTACGCTGTGGACGGGGCTTTTCATCCCAATTTGCACCTACATTGAA 634

Db 127 ArgThrHisLeuTyrValCysGlyThrGlyAlaTyrAsnProIleCysAlaPheIleAsn 146

QY 635 ATTGGACATCATCTGAGGACAATATTTTAAAGCTGGGAAGTACATTTTAAAACCGC 694

Db 147 ArgGlyArgLysAlaGlnAspTyrIlePheTyrLeuGluProAspLysLeuGluSerGly 166

QY 695 CGTGGGAAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAA 754

Db 167 LysGlyLysCysSerTyrAspProLysValAspThrValSerAlaLeuIleAsnGluGlu 186

QY 755 TTATACTCTGGAAGTGCAGCTGATTTTATGGGGCGAGACTTTTGTCTATCTTCGAACTCT 814

Db 187 LeuTyrAlaGlyValTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgThrMet 206

QY 815 GGGCACCACCAATCAGCAGCAGCATGATTCAGGTGCTCAATGATGATCAAG 874

Db 207 GlyLysGlnThrAlaMetArgThrAspGlnTyrAsnSerArgTyrLeuAsnAspProala 226

QY 875 TTCATTAGTCCCACTCATCTCAGAGAGTCACAAATCTGAAGATGACAAAGTATATCTT 934

Db 227 PheValArgAlaGlnLeuIleProAspSerSerGluArgAsnAspAspLysLeuTyrPhe 246

QY 935 TTCTTCGCTGAAATGCAATAGATGGAGAACACCTCTCGAAAAGTACTCACGCTAGAATA 994

Db 247 PhePheArgGluLysSerAlaAspAlaProLeuSer---ProGlyValTyrSerArgIle 265

QY 995 GGTGATATGCAAGAAATGACTTTGGGGGCGACAGAGTCTGTTGAATAAATGGAACA 1054

Db 266 GlyArgIleCysLeuAsnAspAspGlyGlyHisCysCysLeuValAsnLysTyrPheThr 285

QY 1055 TTCTCAAAAGCTCGTCTGATTTGCTCAGTCCAGTCCAAATGGCATTGACACTCATTTT 1114

Db 286 PheLeuLysAlaArgLeuValCysSerValProGlyProAspGlyIleGluThrHisPhe 305

QY 1115 GATGACTGCGAGGATGATCTCTAATGAACCTTTAAAGATCTCTAAATTCAGTGTGATAT 1174

Db 306 AspGluLeuGlnAspValPheIleGlnThrGlnAspThrLysAsnProValIleTyr 325

QY 1175 GGAGTGTTCAGCTTCCAGTAACTTTTCAAGGATCAGCGCTGTGTATGATATGATG 1234

Db 326 AlaValPheSerAlaSerGlySerValPheLysGlySerAlaValCysValTyrSerMet 345

QY 1235 AGTGATGTGAGAAGGTGTTCTTGTCTCATATGCCACAGGATGGAACCAACTATCAA 1294

Db 346 AlaAspIleArgMetValPheAsnArgProPheAlaHisLysGluGlyProAsnTyrGln 365

QY 1295 TGGTGTCTTATCAAGGAAGAGTCCCTATCCAGGCGCAGGAACCTTGTCCAGCAACA 1354

Db 366 TrpMetProTyrThrGlyLysMetProTyrProArgProGlyThrCysProGlyGlyThr 385

QY 1355 TTT---GGTGGTTTGTACTCTACAAAGGACCTTCTGATGATGTTATTAACCTTTCGA 1411

Db 386 PheThrProSerMetLysSerThrLysAspTyrProAspGluValIleAsnPheMetArg 405

QY 1412 AGTCATCCAGCCATGTACAACTCCAGTGTTCCTATGAACAATCGCCCAATAGTGTCAA 1471

Db 406 SerHisProLeuMetTyrHisAlaValTyrProAlaHisArgGlnProLeuValValArg 425

QY 1472 ACGGATGTAATTAATTAATTTACAAATTCGTTAGCCAGGATGCGAGAGATGGA 1531

Db 426 ThrAsnValAsnTyrArgPheThrThrIleAlaValAspGlnValAspAlaAlaAspGly 445

QY 1532 CAGTATGATGTTATGTTTATCGAACAGATCTTGGACCGTCTTAAAGTAGTTTCAATT 1591

Db 446 ArgTyrGluValLeuPheLeuGlyThrAspArgGlyThrValGlnLysValIleValLeu 465

QY 1592 CTAAGAGAGACTTGTGTATGATTTTAGAGAGTCTTCTGGAAGAAATGACAGTTTTCGG 1651

Db 466 ProArgAspAsp---MetGluThrGluGluLeuMetLeuGluGluValPheLys 484

QY 1652 GAACCGACTGTATTTCCAGCAATGGAGCTTTCCTCAATGAGCAACAATATATATGTT 1711

Db 485 ValProAlaProIleLysMetMetThrIleSerLysArgGlnGlnLeuTyrValSer 504


```

QY 1712 TCACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTGATATATTACGGGAAAGCGGT 1771
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505 SerAlaValGlyValThrHisLeuAlaLeuHisArgCysAspValTyrGlyGluAlaCys 524
QY 1772 GCTGAGTGTGCTCGCCGAGACCTTACTGCTTGGTGGTGGTGTCTGTCATGTTCTGCC 1831
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyAlaCysSerArg 544
QY 1832 TATTTTCCCACTGCAAGAGACGACACAGACGACCAAGATATAGAAATGGAGACCACTG 1891
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
545 TyrSerAlaSerSerLysArgArgSerArgArgGlnAspValArgHisGlyAsnProMet 564
QY 1892 ACTCTGTTTACAGCTTACCATGATATACCATGCGCCACGACCCCTGGAAGAGAAATC 1951
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
565 ArgGlnCysArgGly-----TyrAsnSerAsnAlaAsnLysAsnThrValGluAlaVal 582
QY 1952 ATCTATGTTAGAGATAGTAGACATTTTGGATCGATCCGAACTCCGAGTCGACAGAGCG 2011
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
583 GlnTyrGlyValGluGlySerThrAlaPheLeuGluCysGlnProArgSerProGlnAla 602
QY 2012 CTGCTCTATTGCAATTCAGAGCGCAATTAAGAGCGCAAAAGAGATCAGAGTGGAT 2071
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
603 ThrValLysTrpLeuLeuGlnLysAspAsnSerArgArgGlyGluLeuArgThrGlu 622
QY 2072 ---GATCATATCATCAGACAGATCAAGCCCTTCTGCTAGTGTCTACAAGAGGAT 2128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
623 GlyGlyArgAlaLeuArgThrGluGlnGlyLeuLeuLeuArgAlaLeuGlnLeuSerAsp 642
QY 2129 TCAGCAATACCTTCGATCGCGGTGGAACATGGGTTTCATACAACCTCTTCAAGGTA 2188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
643 SerGlyLeuTyrSerCysThrAlaThrGluAsnAsnPheLysHisThrValThrLysVal 662
QY 2189 ACCCTGGAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAAAAGATGATCATGA 2248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 GlnLeuArgValLeuAlaAlaArgAlaValHisAlaValLeuLeuGlnGlyGlu----- 680
QY 2249 GATGCTCTAGACCAAGAAATGTCATAGCATGACACCTPAGCCAGAGGTCTGTGATC 2308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
681 -----LeuProProAlaAlaLeuProGlyAlaProThrProArg-----Tyr 694
QY 2309 AGAGACTTCATGCTCATCAACCCCACTTCAACAGCATGATGATGATGATGATGATGAT 2368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
695 GlnAspLeuLeuGlnLeuThrArgProGluLeuGlyLeuLeuAspGlnTyrCysGln 714
QY 2369 CAAGTTTGGAAAAGGACCGGAAACACAGCTCGGCAAGGCCAGGA---CATACCCCGGG 2425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
715 GlyPheTrp-----ArgProProAlaProGlyProProGluProLeu 728
QY 2426 AACAGTAAACAAATGGAACACTTACAGAAATAGAAAGGTAGAAACAGGAGGACCCAC 2485
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 AlaAlaLeuLysAlaLysGluLeuGlnAspGlnLysLysProArgSerArgArgAsnHis 748

RESULT 8
Q96GX0 PRELIMINARY; PRT; 635 AA.
AC 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC009113; AA090113.1; --
DR InterPro; IPR003599; IG.
DR 240 PropheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 259

```

```

DR InterPro; IPR003006; IG MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 635 AA; 70694 MW; D994099B476B9210 CRC64;

Alignment Scores:
Pred. No.: 3.9e-153 Length: 635
Score: 1905.50 Matches: 359
Percent Similarity: 68.30% Conservative: 102
Best Local Similarity: 53.19% Mismatches: 149
Query Match: 38.92% Indels: 65
DB: 4 Gaps: 10

US-09-774-490-1 (1-2709) x Q96GX0 (1-635)
QY 545 AATTTCATCAGGTACTTAAGGCATATATCAGACTCACTTGTACGCTGTGGACGGG 604
Db 2 AsnPheValLysLeuLeuHisAlaTyrAsnArgThrHisLeuLeuAlaCysGlyThrGly 21
QY 605 GCTTTTCATCAATTTGCACCTACATTCGAAATTTGCACATCATCTCGAGGACAAATATTTT 664
Db 22 AlaPheHisProThrCysAlaPheValGluValGlyHisArgAlaGluGluProValLeu 41
QY 665 AAGCTGGAGAACTCACATTTTAAAGCGCGCTGGGAGAGTCCATATAGCCCTTAAGCTG 724
Db 42 ArgLeuAspProGlyArgGluGluAspGlyLysGlyLysSerProTyrAspProArgHis 61
QY 725 CTGACAGCATCCCTTTTAAATAGATGGAGATTAATCTCTGAACTGCACTGAGCTTTATG 784
Db 62 ArgAlaAlaSerValLeuValGlyGluLeuTyrSerGlyValAlaAlaAspLeuMet 81
QY 785 GGGCGAGACTTGTCTATCTCCGAACCTCTGGGCACACACCACTCAGACAGAGCAG 844
Db 82 GlyArgAspPheThrLeuPheArgSerLeuGlyGlnArgProSerLeuArgThrGluPro 101
QY 845 CATGATTCAGGTGGCTCAATGATCCAAAGTTCATTAGTCCACCTCATCTCAGAGAGT 904
Db 102 HisAspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSer 121
QY 905 GACATCTGAGATGACAAAGTATCTTTTCTCGTGAATGCAATAGATGGAGAA 964
Db 122 GluAsnProAspAspLysIleTyrPhePheArgGluThrAlaValGluAlaAla 141
QY 965 CACTCT---GGAAAAGCTACTCACGCTAGATAGTGCAGATATCAAGAAATGACTTTGGA 1021
Db 142 ProAlaLeuGlyArgLeuSerValSerArgValGlnIleCysArgAsnAspValGly 161
QY 1022 GGGCACAGAGTCTGGTGAATAAATGGACAAACATCTCTCAAAGCTCGTCTGATTGCTCA 1081
Db 162 GlyGlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSer 181
QY 1082 GTGCCAGGTCCAATGCGATTGACACTTATTTGATGAACAGTGCAGAGTATGTTCTTAATG 1141
Db 182 ValProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeu 200
QY 1142 AACTTTAAAGATCCATAAATCCAGTGTATATGAGTGTGTTTACGACTTCCAGTAAACATT 1201
Db 201 SerSerArgAspHisArgThrProLeuLeuTyrAlaValPheSerThrSerSer---Ile 219
QY 1202 TTCAAGGATCAGCCGTGTGTATGTATAGCATGATGTGAGAAGGGTGTCTCTTGGT 1261
Db 220 PheGlnGlySerAlaValCysValTyrSerMetAsnAspValArgAlaPheLeuGly 239
QY 1262 CCATATGCCACAGGATGACCCCACTATCAATGGTGGCTTATCAAGGAAGAGTCCCC 1321
Db 240 PropheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 259

```



```
QY 1322 TATCACGCCGCGAGTCTGCCAGCAAAATTTGGTGGTTTGCACCTCTACAGGAC 1381
Db 260 TyrProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAsp 279
QY 1382 CTTCTCTGATGATGTTATACCTTTGCAAGAGTCATCCAGCCATGTACAAATCCAGTGT 1441
Db 280 PheProAspValIleGlnPheAlaArgAsnHisProLeuMetTyrAsnSerValLeu 299
QY 1442 CTTATGAACAATCGCCCAATAGTATCAAAACGGATGTAATTAATTAACAAAT 1501
Db 300 ProThrGlyArgProLeuPheLeuGlnValGlyAlaAsnTyrPheThrGlnIle 319
QY 1502 GTCTGAGACCGATGATCGACAGATGACAGATGATGATGATGATGATGATGATGAT 1561
Db 320 AlaAlaAspArgValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 339
QY 1562 GTTGGGACCGTCTTAAAGTAGTTTCAATTCCTAAGGACACTGGTATGATTAGAGAG 1621
Db 340 ValGlyThrValLeuLysValIleSerValProLysGlySerArgProSerAlaGluGly 359
QY 1622 GTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGAGCTT 1681
Db 360 LeuLeuLeuGluGluLeuHisValPheGluAspSerAlaAlaValThrSerMetGlnIle 379
QY 1682 TCCACTAACGACACACTATATGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTA 1741
Db 380 SerSerLysArgHisGlnLeuTyrValAlaSerArgSerAlaValAlaGlnIleAlaLeu 399
QY 1742 CACCGGTGTGATATTACCGGAAAGCGTGTGCTGAGTGTGCTGCGCGAGACCTTAC 1801
Db 400 HisArgCysAlaAlaHisGlyArgValCysThrGluCysCysLeuAlaAlaArgAspProTyr 419
QY 1802 TGTGCTGGGATGTTCTGCATGTTCTCGTATTTTCCACATGCGAAGAGCGCAAGA 1861
Db 420 CysAlaTrpAspGlyValAlaCysThrArgPheGlnProSerAlaLysArgPheArg 439
QY 1862 CCACAGATATAGAAATCGAGCCACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
Db 440 ArgGlnAspValArgAsnGlyAspProSerThrLeuCysSer----- 453
QY 1922 CACCATGGCCAGCGCCCTGAA-----GAGAGAATCATCTATGCTGTAGAGAT 1969
Db 454 -----GlyAspSerSerArgProAlaLeuLeuGluHisLysValPheGlyValGluGly 471
QY 1970 AGTAGCACATTTTGAATGCGAGTCGGAAGTCGACAGAGCGCTGCTATTGGCAATTC 2029
Db 472 SerSerAlaPheLeuGluCysGluProArgSerLeuGlnAlaArgValGluTrpThrPhe 491
QY 2030 CAGAGCGGAATGACAGCGCAAGAGAGATCAGATGATGATCATCATCATCAGGACA 2089
Db 492 GlnArgAlaGlyValThrAlaHisThrGlnValLeuAlaGluGluArgThrGluArgThr 511
QY 2090 GATCAAGGCTCTCTGCTAGTCTTACAAAGAGGATTCAGGCAATTCCTCTGCCAT 2149
Db 512 AlaArgGlyLeuLeuLeuLeuArgLeuArgArgAspSerGlyValTyrLeuCysAla 531
QY 2150 GCGGTGGAACATGGGTTCATACAACTCTTCTTAAGTTAACCCCTGGAAGTCAATGACACA 2209
Db 532 AlaValGluGlnGlyPheThrGlnProLeuArgLeuSerLeuHisValLeuSerAla 551
QY 2210 GAGCATTTGGAGAACTTCTTCATAAAGATGATGATGATGATGATGATGATGATGAT 2269
Db 552 ThrGlnAlaGluArgLeu-----AlaArgAlaGluGlu 562
QY 2270 ATGTCCAATPAGCATGACACTAGCCAGAGAGTCTGTGATGAGAGACTTTCATGAGCTCATC 2329
Db 563 AlaAlaProAlaAlaProGlyProLysLeuTyrTyrArgAspPheLeuGlnLeuVal 582
QY 2330 AACACCCCAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2389
Db 582 ----- 582
```

```
QY 2390 AACACAACTCGGCAAGCCAGGACATACCCAGGGAACAGTAACAAATGGAAG----- 2443
Db 593 -----GluProGly---GlyGlyGlySerAlaAsnSerLeuArgMetCys 596
QY 2444 -----CACTTACAAGAAAAATAAGAAAGGAGTAGAAACAGG 2476
Db 597 ArgProGlnProAlaLeuGlnSerLeuProLeuGluSerArgArgLysGlyArgAsnArg 616
QY 2477 AGGACCCAC-----GAATTTGAGAGGCGGCCAGGAGT 2509
Db 617 ArgThrHisAlaProGluProArgAlaGluArgGlyProArgSer 631

RESULT 9
QNS98 PRELIMINARY; PRT; 782 AA.
AC QNS98; Q9H7Q3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Senaphorin sem2 (FLJ00014 protein).
GN SEM2 OR FLJ00014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayaishi A., Kozuma S., Muramatsu M., Miyajima N.,
RT "Human senaphorin."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-782 FROM N.A.
RC TISSUE=Spleen;
RA Chara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
spleen."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029496; BAA98132.1; -
DR EMBL; AK024425; BAB15715.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;

Alignment Scores:
Pred. No.: 1.09e-146 Length: 782
Score: 1830.00 Matches: 366
Percent Similarity: 64.17% Conservative: 139
Best Local Similarity: 46.51% Mismatches: 244
Query Match: 37.38% Indels: 38
DB: 4 Gaps: 13

US-09-774-490-1 (1-2709) x QNS98 (1-782)
QY 221 GTCTGT---CTTCTCTGGGAGTATTACTTACAGAGCAAACTCATCAATGGGAG 277
Db 8 IleCysTrpLeuLeuGlyGlyLeuLeuHisGlyGlySerSerGlyProSerProgly 27
QY 278 AACATGTGCCAAGCTCAATTTATCTTACAAAGAAATGTTGGAATCCAAATGTGATC 337
Db 28 ProSerValProArgLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArgSerAla 47
QY 338 ACTTTCAATGCTTGGCCAAACAGCTCCAGTTTATCATCTCTCTCTTTTGGATGAGAACG 397
Db 338 ACTTTCAATGCTTGGCCAAACAGCTCCAGTTTATCATCTCTCTCTTTTGGATGAGAACG 397
```

Db 48 IlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArg 67
QY 398 AGTAGCTGTATGTTGGAGCAAGAGTACATATTTTCATTCGACCTGGTTAATATC--- 454
Db 68 AspArgLeuPheLeuGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaTrp 87
QY 455 AAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514
Db 88 ProAspProArgGluValLeuTrpProGlnProGlyGlnArgGluGluCysValArg 107
QY 515 GCTGGAAGAACATCTGTAAGAAATGTCTTAATTTTCATCAAGTACTTAAAGGCATATAT 574
Db 108 LysGlyArgAspProLeuThrGluCysAlaAsnPheValArgValLeuGlnProHisAsn 127
QY 575 CAGACTCATTGTGACCTGTGGAACGGGGCTTTTCATCCCAATTTGCACCTACATTTGAA 634
Db 128 ArgThrHisLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuIleThr 147
QY 635 ATTGGACATCCTCGAGACAAATATTTTAAAGCTGGGAACTCACATTTTGAACGGC 694
Db 148 ValGlyHisArgGlyGlu---HisValLeuHisLeuGluProGlySerValGluSerGly 166
QY 695 CTTGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCTTTTAAATAGTGGAA 754
Db 167 ArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPheIleAspGlyGlu 186
QY 755 TTATCTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTGCTATCTTCCGAACTCTT 814
Db 187 LeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetIlePheArgSerGly 206
QY 815 GGGCACCACCCCAATCAGGACAGCAGCATGATTTCCAGTGCTCAATGATCAAG 874
Db 207 GlyProArgProAlaLeuArgSerAspSer---AspGlnSerLeuLeuHisAspProArg 225
QY 875 TTCATTAGTGGCCACCTCATCTCAGAGCTGACATCTCTGAAGATCAGCAAGTATCTTT 934
Db 226 PheValMetAlaAlaArgIleProGluAsnSerAspGlnAsnAspLysValTyrPhe 245
QY 935 TTCCTTCGCGAAAT-----GCAATAGATGGAGAACACTCTCGAAAGACTCACTCAGCT 988
Db 246 PhePheSerGluThrValProSerProAspGly---GlySerAsnHisValThrValSer 264
QY 989 AGAATAGTTCAGATGCAAGAAATGACTTTGGAGGCGCAGAAAGTCTGTGTAATAATGG 1048
Db 265 ArgValGlyArgValCysValAsnAspAlaGlyGlnArgValLeuValAsnLysTrp 284
QY 1049 ACAACATCTCCAAAGCTGCTGTGATTTGCTCAGTCCGAGTCCAAATGCGCATTCACACT 1108
Db 285 SerThrPheLeuLysAlaArgLeuValCysSerValProGlyProGlyGlyAlaGluThr 304
QY 1109 CATTTTGATGAACCTGAGATGATTTCTTAATGAACCTTTAAAGATCTCTAAATAATCCAGTT 1168
Db 305 HisPheAspGlnLeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSerLeuGlu 324
QY 1169 GTATATGAGTGTTCAGATCTCCAGTAACTTTCAAGGATCAGCCGTGTGATGAT 1228
Db 325 ValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCysValTyr 344
QY 1229 AGCATGAGTGTGAGAGGGTGTCTTGTGCTCATATGCCCAAGGATGGACCCAAAC 1288
Db 345 HisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGln 364
QY 1289 TATCAATGGGTCCCTTATCAAGAGAGTCCCTATCCAGGCCAGGAACTTGTCCACGC 1348
Db 365 HisGlnTrpGlyProTyrGlyLysValProPheProArgProGlyValCysProSer 384
QY 1349 AAAACATTTGGT-----GGTTTGTACTTCAAAAGGACCTTCTCGATGATGTT 1396
Db 385 LysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAspGluVal 404
QY 1397 ATACCTTTGCAAGAGTATCCACCATGTACATCCAGTGTTCCTTATGAACATCGC 1456
Db 405 LeuGlnPheAlaArgAlaHisProLeuMetPheTrpProValArgProArgHisGlyArg 424

QY 1457 CCAATAGTATCAAAACGAGTGTAAATTTATCAATTTACAAATTTGCTAGACCCAGTG 1516
Db 425 ProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleValValAspArgVal 444
QY 1517 GATCAGAGATGGACATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCTT 1576
Db 445 GluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThrAspSerGlySerValLeu 464
QY 1577 AAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGA 1636
Db 465 LysValIleAlaLeuGlnAlaGlyGlySerAlaGluProGluGluValValLeuGluGlu 484
QY 1637 ATGACAGTTTTTCGGGAACCGACTCTATTTTCAGCAATGGAGCTTTCCCAATGAGCAA 1696
Db 485 LeuGlnValPheLysValProThrProIleThrGluMetGluIleSerValLysArgGln 504
QY 1697 CAATATATATTTGGTTCACCGCTGGGTGCTCCAGCTCCCTTTTACCGGTGTGATATT 1756
Db 505 MetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnCysGluThr 524
QY 1757 TACGGGAAGCGTGTGCTGAGTGTTCCTCCCGAGAGCCCTTACTGTGCTGGCATGGT 1816
Db 525 TyrGlyThrAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544
QY 1817 TCTGCTATCTTCTCGCTATTTTCCCACT---GCAAGAGAGCGCACAAAGACCAAGATATA 1873
Db 545 AlaserCysThrHisTyrArgProSerLeuGlyLysArgA-pheArgArgGlnAspIle 564
QY 1874 AGAATAGGAGACCCACTGACTGCTTTCAGATTTTACACCATGATTAATCACCATGCCAC 1933
Db 565 ArgHisGlyAsnProAlaLeuGlnCysLeu-----GlyGln 576
QY 1934 AGCCTGGAAGAGAGA-----ATCATCTATGCTGTAGAGAAAT 1969
Db 577 SerGlnGluGluGluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGluHis 596
QY 1970 AGTAGCATTTTTTGGAAATGCGAAGTCCGAGAGCGCTGCTTATTGGCAATTC 2029
Db 597 AsnSerThrPheLeuGluCysLeuProLysSerProGlnAlaAlaValArgTrpLeuLeu 616
QY 2030 CAGAGCGCAATTAAGAGCGGAAAGAGATCAGATGAGTGGATGATCATATCATCAGGACA 2089
Db 617 GlnArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeuHisThr 636
QY 2090 GATCAAGCCTTCTGCTAGTGTACTACACAGAGGATTCAGGCAATTAACCTCTGCCAT 2149
Db 637 GluArgGlyLeuLeuPheArgLeuSerArgPheAspAlaGlyThrTyrThrCysThr 656
QY 2150 GCGGTGGAACATGGTTTCATACAACTCTTAAAGGTAAACCCCTGGAAGTCAATTCACACA 2209
Db 657 ThrLeuGluHisGlyPheSerGlnThrValValArgLeuAlaLeuValIleValAla 676
QY 2210 GAGCATTTGGAAGAACTTCTTATAAAGATGATGAGATGGCTCTTAAAGCCCAAGAA 2269
Db 677 SerGlnLeuAspAsnLeuPheProGluProLysProGluProAlaArgGly 696
QY 2270 ATGTCCAATAGCATGACACCTAGCCAGAGTCTGTTACAGAGATTCATCAGCTCATC 2329
Db 697 GlyLeuAlaSerThrProPro-----LysAlaTrpTyrLysAspIleLeuGlnLeuIle 714
QY 2330 AACCACCCCAATCTCAACAGCATGGATGAGTTCGTGCAACAAAGTTTGGAAAGGACCGCA 2389
Db 715 GlyPheAlaAsnLeuProArgValAspGluTyrCysGluArgValTyrCysArgGlyThr 734
QY 2390 AAACAA-----CGTCGCAAGGCCAGGACATACCCAGGGAACAGT 2431
Db 735 ThrGluCysSerGlyCysPheArgSerArgGlyLysGlnAlaArgGlyLysSer 754
QY 2432 AACAAATGGAAGCACTTACAGAAAAATAGAAAGGTAGAAACAGGAGACCACCAATTT 2491
Db 755 -----TrpAlaGlyLeuGluLeuGlyLysLysMetLysSerArgValHisAlaGluHis 772


```
QY 208 GTTAACAGGATGTC---TGCTTTCTGGGAGTATTACTTACAGCAAGACAACTA 264
Db 22 pMetMetLeuIleMetThrValLeuPheLeuProValThrGluThrSer----- 38
QY 265 TCAGAAATGGGAAGAACAAATGTCACAGGTCGAAATATCTCTACAAAGAAATGTTGAATC 324
Db 39 -----LysGlnAsnIleProArgLeuLysLeuThrTyLysAspLeuLeuLeuSe 55
QY 325 CAACATGTGATCACTTTCAATGGCTTGCCACACAGCTCCAGTTCATCATCTCTTTT 384
Db 55 rAsnThrCysIleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeu 75
QY 385 GCATGAGGACGAGTAGCTGTATGTCGAGCAAGGATCATATTTTCATTCACGCT 444
Db 75 uAspGluGluArgGlyLeuLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuSerLe 95
QY 445 GGTTAATATC---AGGATTTTCAAAGATTGTGGCCAGTATCTTACACCAAGAGAA 501
Db 95 uValAspLeuAsnLysAsnPheLysIleTyTrpProAlaAlaLysGluArgValGl 115
QY 502 TGAATCGAATGGCTGGAAGACATCTGAAAGATGCTGAATTCATCATCAAGTACT 561
Db 115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgValLe 135
QY 562 TAAGGCATATATCACTCACTTGATCGCTGTGGAACGGGGCTTTTCATCAATTTG 621
Db 135 uGlnProTyAsnLysThrHisValTyValCysGlyThrGlyAlaPheHisProLeuCy 155
QY 622 CACCTACATGAAATGGACATCATCTCGAGCAATATTTTAAAGCTGGAGAACTCACA 681
Db 155 sGlyTyrlleAspLeuGlyAlaAsnLysGluGluLeuIlePheLysLeuAspThrHisAs 175
QY 682 TTTTGAAGCGCCGTGGGAAGATCCATATGACCTTAAGCTGCTGACAGCATCTTTT 741
Db 175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMe 195
QY 742 AATAGATGAGAATATATCTCGAACTCAGCTGATTTTATGGCGGAGACTTTGCTAT 801
Db 195 tThrAspGluHisLeuTySerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPh 215
QY 802 CTTCCGAATCTTTGGG-----CACCAACCAATCATCAGCAGCAGCAGCATGTC 852
Db 215 eThrArgSerLeuGlyLeuMetGlnAspHisSerIleArgThrAspIleSerGluHi 235
QY 853 CAGGTGCTCAATGATCCAAAGTTCATTAGTCCACCTCTCAGAGAGGACAATCC 912
Db 235 shistrLeuAsnGlyAlaLysPheIleGlyThrPheProIleProAspThrTyAsnPr 255
QY 913 TGAAGATGACAAAGTATATCTTTCTTCGTAAGAAATGCAATAGATGAGAGACATCTGG 972
Db 255 oAspAspAspLysIleTyPhePhePheArgGluSerSerGlnGluGlySerThrSerAs 275
QY 973 AAAAGCTACTCAGTAGAATAGGTGATGATGCAAGAAATGATCTTTGGAGGGCAGAG 1032
Db 275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSe 295
QY 1033 TCTGTGTAATATGACAACTTCCTCAAGCTGCTGATGTTGCTCAGTCCAGGTC 1092
Db 295 rLeuIleAsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY 1093 AAATGGCATGTGACATCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1152
Db 315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAs 335
QY 1153 TCCTAAATTCAGTTGATATGAGAGTGTTCAGCTTCAGCTTCCAGTAACTTTCAAGGATC 1212
Db 335 pGluArgAsnProValTyThrGlyValPheThrIysThrSerSerIlePheLysGlySe 355
QY 1213 ACCCGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272
Db 355 rAlaValCysValTySerMetAlaAspIleAlaValPheAsnGlyProTyAlaHi 375
QY 1273 CAGGGATGACCCCACTATCAATGGTGTGCTTATCAAGGAAGAGATCCCTATCCAGGCC 1332
```

```
Db 375 sLysGluSerAlaAspHisArgTrpValGlnTyAspGlyArgIleProTyProArgPr 395
QY 1333 AGGAAC 1339
Db 395 oGlyThr 397
RESULT 13
Q9J129 PRELIMINARY; PRT; 296 AA.
ID Q9J129
AC Q9J129;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Brenz Verca M.S., Widmer D.A.J., Wagner G.C., Dreyer J.L.;
RT "New rat semaphorin with closest homology to avian Sema3D (C-
collapsin-2).";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268594; AAF76329.1; -.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00630; Sema; 1.
FT NON_TER 1
FT NON_TER 296
SQ SEQUENCE 296 AA; 33558 MW; C2226212885D044C CRC64;
Alignment Scores:
Pred. No.: 5,08e-83 Length: 296
Score: 1080.00 Matches: 192
Percent Similarity: 81.82% Conservative: 51
Best Local Similarity: 64.65% Mismatches: 52
Query Match: 22.06% Indels: 2
DB: Gaps: 2
US-09-774-490-1 (1-2709) x Q9J129 (1-296)
QY 914 GAAAGTCAAGATATATCTTTCTCGTGAAGAAATGCAATAGATGGAGAACACTCTGGA 973
Db 1 AspAspAspLysIleTyPhePhePheArgGluSerSerGlnGluGlySerThrSerAsp 20
QY 974 AAAGCTACTCAGCTAGATAGTGCAGATATGCAAGATGATCTTTGGAGGGCAGCAAGT 1033
Db 21 ArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSer 40
QY 1034 CTGTGTAATATGAGCAACATCTCTCAAGCTGCTGATGTTGCTCAGTCCAGGTC 1093
Db 41 LeuIleAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerIleProGlySer 60
QY 1094 AATGCATGACATCTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1153
Db 61 AspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAsp 80
QY 1154 CTTAAATCCAGTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1213
Db 81 GluArgAsnProValValTyThrGlyValPheThrThrSerSerIlePheLysGlySer 100
QY 1214 GCGGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
Db 101 AlaValCysValTySerIleAlaAspIleAlaValPheAsnGlyProTyAlaHis 120
QY 1274 AGGGATGACCCCACTATCAATGGTGTGCTTATCAAGGAAGAGTCCCTATCCAGGCCA 1333
Db 121 LysGluSerAlaAspHisArgTrpValGlnTyAspGlyArgIleProTyProArgPr 140
QY 1334 GGAAGTGTCTCCAGCAAAACATTTGGTGGT---TTTGACTCTACAAAGGACCTTCTGAT 1390
```

```
Db 141 GlyThrCysProSerLysThrTyrAspProLeuLeuLysSerThrArgAspPheProAsp 160
QY 1391 GATGTTAAACCTTTGCGCAAGAGTCATCCAGCAGCATGTAACATCCAGTGTTCCTATGAAC 1450
Db 161 AspValIleSerPheIleArgHisProValMetPheLysSerValTyrProValala 180
QY 1451 AATCCCAATAGTGATCAAAACGGATGTAATATCAATTTACACAAATTTGTCGTAGAC 1510
Db 181 GlyAlaProThrPheGlnArgIleAsnValAspTyrArgLeuThrGlnIleValValAsp 200
QY 1511 CGAGTGGATGAGAGATGCGAGATGATGATGTTATGTTATCGGACAGATGTTGGACC 1570
Db 201 HisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAsp***GlyThr 220
QY 1571 GTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTCGTATGATTTTAAAGAGAGTTCGTG 1630
Db 221 ValLeuLysValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeu 239
QY 1631 GAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTACGAAATGGAGCTTTCCACTAAG 1690
Db 240 GluGluLeuGlnValPheLysHisProThrAlaIleLeuAsnMetGluLeuSerLeuLys 259
QY 1691 CAGCAACATATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGT 1750
Db 260 GlnGlnGlnLeuTyrValGlySerTrpAspGlyLeuValGlnLeuSerLeuHisArgCys 279
QY 1751 GATATTTACGGGAACCGTGTGCTGAGTGTGCTGCTGCGCCGAGAGCCCTTAC 1801
Db 280 AspThrTyrGlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyr 296
```

RESULT 14

```
Q8BJC1
ID Q8BJC1 PRELIMINARY; PRT; 799 AA.
```

```
AC Q8BJC1;
```

```
DT 01-MAR-2003 (TReMBLrel. 23, Created)
```

```
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
```

```
DE Senaphorin 4D precursor (Fragment).
```

```
OS Mus musculus (Mouse).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
NCBI_TaxId=10090;
```

```
SEQUENCE FROM N.A.
```

```
RC STRAIN=NOD; TISSUE=Thymus;
```

```
RX MEDLINE=22354683; PubMed=12466851;
```

```
RA The FANTOM Consortium,
```

```
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
```

```
RT "Analysis of the mouse transcriptome based on functional annotation of
```

```
RL Nature 420:563-573(2002).
```

```
DR EMBL; AK088653; BAC40480.1; --
```

```
FT NON TER
```

```
SQ SEQUENCE 799 AA; 88814 MW; 415D3B687150A59A CRC64;
```

```
Alignment Scores:
```

```
Pred. No.: 4.18e-74 Length: 799
```

```
Score: 976.50 Matches: 253
```

```
Percent Similarity: 53.53% Conservative: 111
```

```
Best Local Similarity: 37.21% Mismatches: 238
```

```
Query Match: 19.94% Indels: 78
```

```
DB: 11 Gaps: 27
```

```
US-09-774-490-1 (1-2709) x Q8BJC1 (1-799)
```

```
QY 407 TATGTTGGCAAGATGACATATTTTCATTCGACCTGGTAAATATCAAGATTTTCAA 466
```

```
Db 1 TyrValGlyAlaArgGluAlaValPheAlaValAsnAlaLeuAsnIleSerGluLysGln 20
```

```
QY 467 ---AGATTGTGTCAGTATCTTACACAGAGAGATGATGCAAGTGGGCTGGAAA 523
```

```
Db 21 HisGluValTyrTrpLysValSerGluAspLysSerLysCysAlaGluLysGlyLys 40
```

```
QY 524 GACATCTCTGAAGAATGTGCTTAATTTTCATCAAGGTACTTAAGGCATATAATACAGACTCAC 583
Db 41 SerLysGlnThrGluCysLeuAsnTyrIleArgValLeuGlnProLeuSerArgThrSer 60
QY 584 TTGTACGCTGTGAACGGGGCTTTTCATCAATTTCCACCTACATTTGAATTTGGACAT 643
Db 61 LeuTyrValCysGlyThrAsnAlaPheGlnProThrCysAspHisLeuAsnLeu----- 78
QY 644 CATCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACCGCGCTGGGAAG 703
Db 79 -----ThrSerPheLysPheLeuGlyLysSer---GluAspGlyLysGlyArg 93
QY 704 AGTCCATATACCTTAAGCTGCTGACGACATCCCTTTTAAATAGATGGAGAATTATACCT 763
Db 94 CysProPheAspProAlaHisSerTyrThrSerValMetValGlyGlyGluLeuTyrSer 113
QY 764 GGAAGCTGAGCTGATTTATGGGGCGAGACTTCTCTATCTCCGAACTCTTGGGGCACCAC 823
Db 114 GlyThrSerTyrAsnPheLeuGlySerGluProIleLeuSerArgAsn---SerSerHis 132
QY 824 CACCAATCAGACAGACAGCAGCATGATCCAGGTGGCTCAATGATCCAAAGTTCAATTAGT 883
Db 133 SerProLeuArgThrGlu---TyrAlaIleProIleLeuAsnGluProSerPheValPhe 151
QY 884 GCCCAGCTCATCTCAGAGAGT---GACAAATCCT-----GAAGATGACAAAGTATACCTT 934
Db 152 AlaAspValIleGlnLysSerProAspGlyProGluGlyGluAspAspLysValTyrPhe 171
QY 935 TTCTTCGCTGAAATGCAATAGATAGGAGAACACTCTCGAAAGAGTACTCAGCTAGATA 994
Db 172 PhePheThrGluValSerValGluTyrGluPheValPheLysLeuMetIleProArgVal 191
QY 995 GGTGAGATATGCAAGATGACTTTTGGGGGACAGAGTCTGCTGAATAATGACACACA 1054
Db 192 AlaArgValCysLysGlyAspGlnGlyLeuArgThrLeuGlnLysLysTrpThrSer 211
QY 1055 TTCTTCAAGCTCGTCTGATTTGCTGATGTCAGGTGCGAGTCCAAATGGCATGACATCATTT 1114
Db 212 PheLeuLysAlaArgLeuIleCysSerLysProAspSerGlyLeuVal-----Phe 228
QY 1115 GATGAGCTGAGGATGATTTCTTAATGAACTTTTAAAGATCTCAAAATCCAGTTGTATAT 1174
Db 229 AsnIleLeuGlnAspValPheValLeuArgAlaProGlyLeuLysGluProValPheTyr 248
QY 1175 GGAGTGTTCAGCTCCAGTAACATTTTCAAGGATGACCGGTGCTGATGATAGCATG 1234
Db 249 AlaValPheThrProGlnLeuAsnAsnValGlyLeuSerAlaValCysAlaTyrThrLeu 268
QY 1235 AGTATGTGAGAGGGGTGTTTC---CTTGGTCCATAT-----GCCACAGGATGGA 1282
Db 269 AlaThrValGluAlaValPheSerArgGlyLysTyrMetGlnSerAlaThrValGluGln 288
QY 1283 CCCAACTATCAATGGGTGCTTATCAAGGAAGATGTCCTTATCCAGCCGACGAGCTTGT 1342
Db 289 SerHisThrLysTrpValArgTyrAsnGlyProValProThrProArgProGlyAlaCys 308
QY 1343 -----CCCACAAAACATTTGGTGGTTTGGCTCTACAAAGGACCTTCTGTATGATGTT 1396
Db 309 IleAspSerGluAlaArgAlaAlaAsnTyrThrSerSerLeuAsnLeuProAspLysThr 328
QY 1397 ATAACCTTTGCAAGAGTATCCAGCATGTAACATCCAGTGTTCCTTATCAACAAATCGC 1456
Db 329 LeuGlnPheValLysAspHisProLeuMetAspAspSerValThrProIleAspAsnArg 348
QY 1457 CCAATAGTATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGAGACCGAGTG 1516
Db 349 ProlLysLeuIleLysLysAspValAsnTyr-----ThrGlnIleValValAspArgThr 366
QY 1517 GATCAGAAAGATGACAG-----TATGATGTTATGTTTATCGGAACAGATGTTGGACCGTT 1573
Db 367 GlnAlaLeuAspGlyThrPheTyrAspValMetPheIleSerThrAspArgGlyAlaLeu 386
```



```
Db 378 CysTyrAlaGluGlnValValAlaAlaArgValCysGlyCysGlyAspMetGlyGly 397
QY 1025 CACAGAAGTCGTGGTAATAAATGGAACATTCCTCAAGCTCGTCTGATTTGCTCAGTG 1084
Db 398 AlaArgThrLeuGlnArgLysTrpThrPheLeuLysAlaArgLeuAlaCysSerAla 417
QY 1085 CCAGTCCCAATGGCATTGACACTCATTTTGTATGAACTGCAGGATGTATTCCTAATGAAC 1144
Db 418 Pro-----AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGln 434
QY 1145 TTTAAAGATCCTAAATAATCCAGTGTGTATATGAGGTGTTTACGACTTCCAGTAACATTTTC 1204
Db 435 AspThrSerTrpHisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMet 454
QY 1205 AAGGATCAGCGCTGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
Db 455 TyrLeuSerAlaIleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyPro 474
QY 1265 TATGCCACAGGATGGACCAACTATCAATGGTGGCTTATCAAGGAGAGTCCCTTAT 1324
Db 475 TyrLysGluTyrHisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSer 494
QY 1325 CCACCGCCAGGAACTTGTCCCAAGCAAAACATTT-----GGTGGTTTTGACTCTACAAAG 1378
Db 495 ProArgProGlySerCysIleAsnAsnTrpHisArgArgHisGlyTyrThrSerSerLeu 514
QY 1379 GACCTTCCTGATGATGATTTATACCTTTGCAAGAGTCATCCAGCCATGTACAAATCCAGTG 1438
Db 515 GluLeuProAspAsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnVal 534
QY 1439 TTTCTTATGACAATCGCCCAATAGTATGATCAAAACGGATGTAAATTAATTAATTAACAA 1498
Db 535 GlyProArgTrpSerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHis 552
QY 1499 ATTGCGTAGACCGAGTGCATCAGACAGATGA-----CAGTATGATGTTATGTTATTCGA 1555
Db 553 LeuValAlaAspArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGly 572
QY 1556 ACAGATGTTGGACCGTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTA 1615
Db 573 ThrGlyAspGlyTrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis--- 589
QY 1616 GNAGAGTTTCCTCGGAAGAATGACAGTTT-----CGGGAACCGACTGCTATTTTCAGCA 1672
Db 590 -----LeuIleGluGluLeuGlnPheAspGlnGluPro-----MetArgSer 604
QY 1673 ATGGAGCTTTCCTCAAGCAGCAACATATATATTTGTTCAACGGCTGGGGTTCGCCAG 1732
Db 605 LeuValLeuSerGlnSerLysLysLeuLeuPheAlaGlySerArgSerGlnLeuValGln 624
QY 1733 CTCCCTTTACACCGGTGTGATATTTACGGGAAGCGTGTGTGAGTGTGCTCGCCCGA 1792
Db 625 LeuProValAlaAspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArg 643
QY 1793 GACCTTACTGTGCTTGGATGGTTCGTCGATGTTCTCGCTATTTTCCCACTGCAAGAGA 1852
Db 644 AspProTyrCysAlaTrpSerValAsn----- 652
QY 1853 CGCACAGACGA-----CAAGATATA 1873
Db 653 ---ThrSerArgCysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisVal 671
QY 1874 AGAATGAGAGCCCACTGACTCCTGTTTCCAGACTTACACCATGATAATCACCATGGCCAC 1933
Db 672 MetThrSerAspThrSerGlyIleCys-----AsnLeuArgGlySer 685
QY 1934 AGC-----CCTGAAGAGAGAAATCATCTATGTTGTAGAGATAGTACACATTTTGT 1984
Db 686 LysLysValAlaArgProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeu 705
QY 1985 GAATCAGTCCGAGTCCGAGAGCGCTGCTATTTGCAATTCAGAGCGGAAT--- 2041
Db 706 ProCysHisLeuSerSerAsnLeuAlaHisAlaArgTrpThrPheGlyGlyArgAspLeu 725
```

```
QY 2042 ---GAAGAGCGAAAGAGAGATCAGATGGATGATCATATCATCAGGACAGATCAAGGC 2098
Db 726 ProAlaGluGlnProGlySerPheLeuTyrAspAlaArgLeu-----GlnAla 741
QY 2099 CTTCGTCTACGTAGTCTTCAACAGAGGATTCAGGCAATTACCTCTGCCATCGGTGGAA 2158
Db 742 LeuValValMetAlaAlaGlnProArgHisAlaGlyAlaTyrHisCysPheSerGluGlu 761
QY 2159 CATGGG 2164
Db 762 GlnGly 763
```

Search completed: August 3, 2003, 10:01:25
Job time : 166 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus n2p model

Run on: August 3, 2003, 10:02:39 ; Search time 63 Seconds
(without alignments)
10213.340 Million cells

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatcttttattttatcgaqg.....agqctttttttcctaataacc 2709

Scoring table:

Xgapop	10.0	, Xgapext	0.5
Ygapop	10.0	, Ygapext	0.5
Fgapop	6.0	, Fgapext	7.0
Delop	6.0	, Delext	7.0

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 903798

Minimum DB seq length: 0

Maximum DB seq	length: 0
Maximum DB seq	length: 2000000000

post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Command line parameters:

```

MODEL=frame+rnp_model_DEV=xlh
-Q/cgmn2/1/USTO_spool/US09774490/runat_03082003_095418_25421/app_query.fasta_1.2887
-Db=PublishedApplicationsAA-QFMT=fastan-SUFFIX=rapb-MINMATCH=0.1
-LOOPEL=0-LOOPEXT=0-UNITS=bits-START=1-END=1-PATCX=blslogm2
-TRANS=human40.cdi-LIST=45-DLOCALIGN=200-THR_SCORE=1-THR_MAX=100
-THR_MIN=0-ALIGN=15-MODE=LOCAL-OUTFWT=ptc-NORM=ext-HEAPSIZE=500-MINLEN=0
-MAXLEN=200000000-USER=US09774490@CGN_1_1_85@runat_03082003_095418_25421
-NCPU=6-ICPU=3-NO MMAP-LARGQUERY-NEG_SCORES=0-WAIT-DSPBLOCK=100
-LOGLOG-DEV TIMEOUT=120-WARN TIMEOUT=30-THREADS=1-XGAPOP=1-XGAPEXT=0.5
-XGAPOP=6-FGAPEXT=7-XGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

```

Database : Published Applications AA.*

```

Database :
Published_Aplications_AAI:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	100	100	100	100	100	100

ALIGNMENTS

RESULT 1

US-10-097-340-284
; Sequence 284, Application US/10097340
; Publication No. US20030087250A1

: GENERAL INFORMATION:

APPLICANT: JOHN MONAHAN

APPLICANT: Manjula GANNAVARAPU

APPLICANT: Sebastian Hoersch

; APPLICANT: SEDBCLIAN HOERSCH
: APPLICANT: SHUBHANGI KAMATKAR

APPLICANT: SHUBHANGI K

APPLICANT: Steve G. KO
APPLICANT: Rachel E. M

APPLICANT: Rachel E. MEYERS

APPLICANT: MICHAEL MORRIS

APPLICANT: Peter OLAND

; APPLICANT: Ami SEN

APPLICANT: Peter VEIBY

APPLICANT: Gordon B. Mills

APPLICANT: Robert C. B.

APPLICANT: Karen LU

APPLICANT: Rosemarie S

; APPLICANT: Xumei ZHAO

APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,

TITLE OF INVENTION: ASH

```

; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

```

Alignment Scores:

```

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

```

QY 200 ATGGGCTGTAACTAGGATGTCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAAGCAATGTGCCAAGCTGAAATATCTTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAACTATGTATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnValIlePheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATCAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGACA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTCGAAAGACATCTCGAAAGATGTCTTAATTTTCATCAAGGTA 559
DB 101 AspGluCysIleValGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCCTTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAATTTGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACGGCGTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAAATATATCTCTGGAAGTGCAGCTGATTTTATGGGGCAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

```

```

QY 800 ATCTTCGAACTCTTGGSCACACCCCAATCAGGACAGAGCAGCATGATTCTCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATCAATAGATGGAGAACACTCTGGAAGACT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTACAGATATGCAAGATGACCTTTGGAGGSCACAGAGTCTGGTG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGCAACAATCTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTTGATGAACCTGAGGATGTATCTTAATGAACCTTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGGAGTCTTTACGACTTCCAGTACATTTTCAAGGGATCAGCGTG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGAGAGGGTGTCTTCTGTGTCATATGCCACAGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGAGAGAGTCCCTATATCCAGGCCAGAACT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCGCCAGCAAAACATTTGGTGTTCACACTCAAGAGGACCTTCTGTATGATGCTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTGTCTTCTATGAACATCCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAAATTCATTTTACAAATTTCTGTAGACCGAGTGGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGCACAGTATGATGTTATCGAACAGATGTTGGACCGTCTCTTAAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 CTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGGTTCTCTCGAAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTACCAATGGAGCTTCCACTAAGCAGCAACAA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGluLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGTTCCTCCCGGAGACCCCTTACTGTGCTTGGAGTGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGAGCGCACAGCAGACAGATATAGAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560

```

QY 1880 GGAGACCCACTGACTGCTTCTGAGCTTACACCATGATAATCACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGCACATTTTGGAAATGCAGTCCGAAG 1999
Db 581 GluGluArgLleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProlys 600
QY 2000 TCGAGAGAGCGCTGCTTATTGGCAATTCAGAGCGGAAATCAAGAGCGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATCATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACATGGTTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheLeuGlnThrLeu 660
QY 2180 CTTAAGTTAACCTGGAAGTCAATTCAGACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCGTGTACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACACAGATGATGAG 2359
Db 701 ValTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAAGTTTGGAAAAGGACCGAAAACACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAGACATTCAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGGAATTTGAGAGGCGCCAGGAGGTGC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

US-10-262-538-10
; Sequence 10; Application US/10262538
; Publication No. US2003011324A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

Alignment Scores:

Pred. No.:	0	Length:	771
Score:	4201.00	Matches:	771
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.80%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTCGTCTTTCTGGGAGTATTACTTACAGCAAGACGA 259

Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGGAACAACAATGTGCCAAGGCTGAAATTTATCTCAACAAGAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGAACCGAGTAGGCTGTATCTTGGAGCAAAAGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCTGGAAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheLysVal 120
QY 560 CTTAAGGCATATATCAGACTCATTGTACCCCTGTGGAAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCAGAGACAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisIleProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTGAAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTCCAGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTCTGGGCACCCACCACCAATCAGGACAGAGCAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisIleHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLeuSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATACTTTTCTCCGTGAAAATGCCAATAGATGGAGAACACTCTGGAAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACGCTGAATAGTTCAGATATGCAAGAATCACTTTGGAGGGCCAGAAAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATTCCTCAAAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCTCAAAATGCG 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTATGAATCTGACGAGTGTATCTTAATGAACCTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTGTATATGAGTGTATTACACTTCCAGTAAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTATGAGAAGGCTGTCTCTGCTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATCGGTGCTTATCAAGGAAGAGTCCCTTATCCACGCCCAAGAACT 1339

Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGGACTCTACAAAGACCTTCCTCATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTCAATCCAGTGTTCCTATGAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTACACAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGGACAGATGATGTTATGTTATTCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACATGTTTTCGGAACCGACTGCTATTTCCAGCAATGGAGCTTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGATATTATAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTCTCAGTGTTCCTCGCCGAGACCTTACTGTCTGGATGGTCTCT 1819
Db 521 GlyValAlaCysAlaGluCysLeuAlaArgAspProTyrCysAlaIleAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGACGCAAGACGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTCCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATCGTGTAGAGATAGTAGCACATTTTGGAAATGCAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCGCTGCTCTATTGGCAATTCAGAGCGCAAAATCAAGAGCAAAAGAAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTCGATCATATCATCAGCAGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGATTTCAGGCAATTTACCTCTCCCATGGGTGGAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCATTGACACAGACGATTTGGAAGAACTTCTTCATAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCTAATAGATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGTACAGACTTCATGACGCTCATCAACCCCACTCTCAACACGATGATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTGGAAAGGGACCGAAACAAACGTCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

QY 2420 CCAGGACACTACAAATGGAACGACTTCAAGAAATAAGAAAGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGGCACCAGGAGTGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3

US-09-946-374-310
; Sequence 310, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808

;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099815
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100388
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100584
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100661
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100662
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100664
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100683
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100684
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100710
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100711
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100849
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100930
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101014
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101068
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101071
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101279
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: 60/101471
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101472
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101474
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101475
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101476
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101479
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101743
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101915
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/102207
;; PRIOR FILING DATE: 1998-09-29

;; PRIOR APPLICATION NUMBER: 60/102240
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102307
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102330
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102331
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102484
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102487
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102571
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102684
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102687
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102965
;; PRIOR FILING DATE: 1998-10-02
;; PRIOR APPLICATION NUMBER: 60/103258
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103314
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103315
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103328
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103395
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103396
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103401
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103633
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/104257
;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Alignment Scores:

Pred. No.: 1,05e-199
Score: 2245.50
Percent Similarity: 73.67%
Best Local Similarity: 54.92%
Query Match: 45.86%
DB: 11
Length: 777
Matches: 413
Conservative: 141
Mismatch: 177
Indels: 21
Gaps: 9

QY 2420 CCAGGGAACAGTAACTGGAAGCACTTACAGAAAATAAGAAAGGTGAAACAGGAGG 2479
Db 748 -----ProlystirpLySHisMetGlnGluMetLySLysArgaenArgAG 763
QY 2480 ACCCAC---GAATTTCAGAGGCGACCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 4

US-10-052-586-348
; Sequence 348, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22

Db 379 AlaAspHisArgTrpValGlnThrAspGlyArgGlnProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTGGTGT---TTTGACTCTACAAAGAGACCTTCCTGATGATGTATA 1399
Db 399 ProSerLysThrTyrAspProLeuLeuLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTACCCAGCCATGATCAATCAGTGTTCCTATGACAAATGCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACCGATTAATATCAATTTACACAATTTGCTAGACCGAGTGGAT 1519
Db 439 ThrPheLysArgIleasnValAspTyrArgLeuThrGlnIleValValAspHisValle 458
QY 1520 GCAGAGATGACAGATGATGATTTATCGGAACAGATGTTGGACCGTCTCTTAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 CTAGTTTCATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGATG 1639
Db 479 ValValSerIleSerLysGlnLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTCTATTTCAGCAATGGAGCTTCCACAAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTGCCAGCTCCCTTTACACGGTGTGATATTTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGCTGCTGGATGGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTCCTCCACTCAAGAGACCCACAGACAGATATAGAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCTGTTTACAGATTTACACCATGATATACACCTGCGCCAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GNAGAGAGATCATCTATGCTGATAGATATAGTACACATTTTGGATGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGCGCTGCTGTTTGGCAATTCAGAGCGCAATAGAGCGCAAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAGTCTCAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGlnTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAGGATTCGGCAATTTACTCTGCCATCGGTGGAAACATGGGTTCATCAAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAATCTATTCACACAGCATTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGGAGATGGCTTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACACAGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACACACCTCGCCAAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetIrrPheArgGlnLysArgGlnArgGlnArgGlnLysGlyGly----- 747

QY 2420 CCAGGAACAGTAACAATGAAGCACCTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGGCCACCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
RESULT 5
US-10-174-590-348
; Sequence 348, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; PRIORITY FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-348
Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
Gaps: 9
US-09-774-490-1 (1-2709) x US-10-174-590-348 (1-777)
QY 275 AAGAACAAATGTCCAAAGGCTGAAATTTATCTTACAAAGAAATGTTGGAAATCAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCTCTCTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGGAGTAGTGGTGTATGTTGGCAAGCAAGATACATATTTTCATTCGACCTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuLeuAspLeu 98
QY 455 ---AAGATTTCAAAAGATTGTGTGCCAGATCTTACACCAAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTCGAAAGACATCTCGAAGAAATGTCTTAATTTTCATCAAGGTACTTAAGGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCATTTGATGCTGTGGAACCGGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGGACATCATCTCTAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAC 691

Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCGGTGGGAAGAGTCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATATATCTCTGGAACTCCACTGATTTATGGGCGAGACTTTGCTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCAACCAATCAGACAGACGATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleAspHisIleAspHisIleAspHisIleAspHisIle 238
QY 863 AATGATCCAAATTCATTAGTCCCACTCATCTCAGAGAGTGAGCAATCCTGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAAGTATATCTTTCTTCCTCGTCAAAATGCAATAGATGGAGACACTCTGGAAGCTACT 982
Db 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CACGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGACACAGAACTCTGCTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGACAAATCTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCATT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTGATCAACTGAGATGTATTCCTAATGCACTTTAAAGATCCTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTAT 1222
Db 339 ProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTATCATGGTGGCTTATCAAGAGAGTCCCTATCCAGCCGAGGACTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTGATGATGTATA 1399
Db 399 ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGATCATCCAGCATGATCAATCCAGTGTTCCTATGAACCAATCCGCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTATCAAAACGGATGAATATATCAATTTACAAATTTGCTAGACCGATGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAGATGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGAACCGACTCTATTTCAGCAATGGACCTTCCACTAGCAGCAACAA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLysGlnGln 517
QY 1700 CTATATATGTTTCAACGGCTGGGTGCGGAGTTCCTTACCGGTGTATATTTAC 1759

Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTCTGAGTGTGCTCGCCGAGACCTTACTGCTGCTGGATGGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTGCTCTATTTTCCCACTGCAAGAGAGCGCACAGACGCAACAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCTACTGTTTCACACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGTTGAGTATAGTACACATTTTGGATGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGAGCGTGTCTATTGGCAATTTCAGAGGCGAAATGAAGCGCAAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATCATCAGGACAGATCAAGGCTTCTGCTAGTAGTCTACA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGATTCCAGCAATTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2179
Db 657 LysLysAspSerGlyMetTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAGTCAATTCAGACACGATTTGGAAGAACTTCTTCTATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGCTCAATGATGATGATGATGATGATGAT 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGAGTCTCATGCTCAGCTCATCAACACCCCAATCTCAACACCATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTCAACAAGTTTGGAAAGGACCGCAAAACAACTCGGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACAGTAAACAATGGAAGCTTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnMetLysLysLysLysLysLysLysLysLys 763
QY 2480 ACCAC---GAATTTGAGAGGCGCCAGGAGTCTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 6

US-10-176-758-348
; Sequence 348, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C104

Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
Qy 2180 CTTAGGTAACTGGAAGCTATTGACACAGAGCATTTGGAGAACTTCTTCAATAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
Qy 2240 GATGATGGAGATGGCTCTAAGCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlnValLys-----AspLeuLeuAlaGluSerArg 710
Qy 2300 GTCTGTACAGACTTCATGAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe-----SerLeuAspGln 729
Qy 2360 TTCTGTCAACAGATTGGAAAGGACCGAAACAAACACGTGCGCAAGGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
Qy 2420 CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATAACAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
Qy 2480 ACCAC- --GAATTTGAGAGGCGCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 7

US-10-175-737-348
; Sequence 348, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; Prior Filing Date: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-175-737-348 (1-777)

Qy 275 AAGAACAAATGTGCAAGGTGAAATATCTCAAAAGAAATGTGGAATCCAAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
Qy 335 ATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCTTTTGGATGGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78

Qy 395 CGAGTAGGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTCGACCTCGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
Qy 455 ---AAGATTTTCAAAGATTGTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
Qy 512 TGGCTGGAAGACATCTCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTAAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
Qy 572 AATCAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCATCCAAATTTGACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
Qy 632 GAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAANAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
Qy 692 GGGCTGGAGAGCTCCATATGACCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
Qy 752 GAATTTACTCTGGAACCTGACAGCTGATTTTATGGGCGAGACTTTTGCTTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
Qy 812 CTTGGG-----CACCACCACCATCAGGACAGCAGCATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleArgThrAspIleSerGluHisTyrTrpLeu 238
Qy 863 AATGATCCAAAGTTTCATTAGTCCACCTCATCTCAGAGAGTGACAATCTCTGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrThrAsnProAspAsp 258
Qy 923 AAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACT 982
Db 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
Qy 983 CAGCTAGATAGTTCAGATATGCAAGATGACATTTGGAGGCGCAGAGTCTGGTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuLeuAsn 298
Qy 1043 AAATGGCAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTGCCAAATGGCAT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
Qy 1103 GACACTCATTTTGATGAACCTCAGAGATGATTTCTTAATGAATTTTAAAGATCTCTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
Qy 1163 CCAGTTGTATATGAGTGTTTACGACTTCCAGTAACTTTTCAAGGGATCAGCCGTGCT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
Qy 1223 ATGTATAGCATGATGATGCGAAGGCTGTTCTTGTGCTCATATGCCACAGAGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
Qy 1283 CCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
Qy 1343 CCCAGCAAAACATTTGGTGGT---TTTACTCTCAAAAGGACCTTCTCTGATGATGATATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle 418
Qy 1400 ACCTTTTGAAGAGTCCAGCCCATGTACAATCCAGTGTGTTTCTTATCAACAATGCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
Qy 1460 ATAGTGATCAAAACGGATGTAAATTTACAAATTTACAAATTTGCTGAGACCGAGTGAT 1519


```
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGCTTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTGCTGATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGlyLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGTGTGATATTATC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGCGTGTGCTGAGTGTGCTGCTGCCGAGACCTTACTGTGTTGGATGTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAen 557
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGACAGCAGCAGCAGCAAGATATAAGAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCAGTCTTACAGCTTACACCATGATATACCATCGGCCAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGCTAGAGATAGTAGACATTTTTCGAATTCAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGCGCTGCTGTTATGGCAATTCAGAGCGCAATAGACAGGCAAAAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGlu 636
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTCTAGTGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTCAGCAATACCTCTGCCATCGGTGGACATCGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAACTCGAAGTCAATTCACACAGCAGCATTTGGAAGAATCTTCTCATNAAG 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGGAGTGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGAGCTCATCAACACCCCAATCTCAACACGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGAAGAAGGACGCAAAACACGTCGGCAAGGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGAACAGTAAACAAATGGAAGCACATTACAGAAATAAGAGAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCAC---GAATTTGAGAGGCCACCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
RESULT 8
US-10-173-706-348
```

```
; Sequence 348, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservatives: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-173-706-348 (1-777)
QY 275 AAGAACATGTGCCAAGGCTGAATATTCCTACAAAGAAATGTTGGAATCCAAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGCTGTGGCCACAGCTCCAGTTCATCATCATCTCTCTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGGAGTAGCTGTATGTTGGAGCAAGCATCATATTTTCATTCGACCTCGCTGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGGCTGAAAGACATCTCGAAAGATGTCTAATTTTCATCAAGGTACTTAAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTGTAGCTGTGGACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTCGACATCATCTCGAGGACAAATTTTAAAGCTGGAGAACTCATTGTAATAAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTAATCTCGAACTGAGCTGATTTTATGCGGCGAGACTTTGCTATCTCTCCGAACT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
```

812 CTTGGG-----CACCACCAACCAATCAGGACAGACGACGATGATTCCAGGTGGCTC 862
Db |||||
219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrIlePleu 238
Qy |||||
863 AATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCTCGAAGATGAC 922
Db |||||
239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrAsnProAspAspAsp 258
Qy |||||
923 AAGATATATCTTTCTTCCTCGTGAATGCAATAGATGAGAGACACTCTGGAAGAGTACT 982
Db |||||
259 LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
Qy |||||
983 CAGCTAGATAAGTTCAGATATGCAAGATGCTTTGGAGGACACAGAGTCTGGTGAAT 1042
Db |||||
279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
Qy |||||
1043 AATGAGACAACTTCCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCAAT 1102
Db |||||
299 LysTyrThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
Qy |||||
1103 GACACTCAATTTGATGAACTGAGGATGATTCTTAATGAACCTTTAAAGATCCTAAAT 1162
Db |||||
319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
Qy |||||
1163 CAGTTGTATATGGAGTGTTCAGACTCCAGTAACTTTTCAAGGATCAGCGGTGT 1222
Db |||||
339 ProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySerAlaValCys 358
Qy |||||
1223 ATGTATAGCATGATGATGAGAGGGTGTCTTCTGTCCTATGTCCTCAATGAGGATGGA 1282
Db |||||
359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
Qy |||||
1283 CCAACTATCAATGGTGCTTATCAAGAGAGTCCCTATCCAGGCGCAGCACTGT 1342
Db |||||
379 AlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
Qy |||||
1343 CCAGCAAAACATTTGGTGT---TTTGACTCTCAAAAGGACTTCTCTGATGATGTATA 1399
Db |||||
399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
Qy |||||
1400 ACCTTTGCAAGATGATCCAGCATGTACATCCAGTGTTCCTATGAAACATCGCCCA 1459
Db |||||
419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
Qy |||||
1460 ATAGTATCAAAAGGATGATTAATATCAATTAACAATTTGCTAGACCGAGTGTAT 1519
Db |||||
439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
Qy |||||
1520 CGAAGATGAGCAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAA 1579
Db |||||
459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
Qy |||||
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCCTGCGAAGAAATG 1639
Db |||||
479 ValValSerIleSerLysGluLysTyr---AsnMetGluValValLeuGluLeu 497
Qy |||||
1640 ACAGTTTTCGGAACCGACTGCTATTTACGACATGAGGACTTCCACTAAGCAGCAGCAA 1699
Db |||||
498 GlnIlePheLysHisSerIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
Qy |||||
1700 CTATATATTTGTTCAACGCTGGGTTCGCCAGTCCCTTTACACGGGTGATATTTAC 1759
Db |||||
518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
Qy |||||
1760 GGAAGACGCTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTTGGATGTTCT 1819
Db |||||
538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyAsn 557
Qy |||||
1820 GATGTTCTCGCTATTTTCCACTGCAAGAGCAGCAGCAGCAGCAGATATAAGAAAT 1879
Db |||||
558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
Qy |||||
1880 GGAGACCCACTGACTCTCACTGTTTCAGACTTTACACCATGATTAATCACCATGGCCACGCCCT 1939

578 GlyAspProIleThrGlnCysTyrAspIleGluAspSerIleSerHis---GluThrAla 596
Qy |||||
1940 GAAAGAGAAATCATCTATGTTGTAGAGATAGTACACATTTTGGATGCGAGTCCGAG 1999
Db |||||
597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
Qy |||||
2000 TCGCAGAGAGCGTGTCTATTGGCAATTTCAGAGGCGCAATTCAGAGAGCGCAAAAGAG 2059
Db |||||
617 SerGlnGlnAlaThrIleLysTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
Qy |||||
2060 ATCAGAGTGGATGATCATCATCATCAGGACAGACATCAAGCCCTTCTGCTAGTGTACAA 2119
Db |||||
637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
Qy |||||
2120 CAGAAGATTTCAGCAATTAATCTCGCATCGGTGGACATGGTTCATACAACTCTT 2179
Db |||||
657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
Qy |||||
2180 CTTAAGGTAAACCTGGAAGTCAATGACACAGACATTTTGGAGAACTTCTTCATAAAGAT 2239
Db |||||
677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
Qy |||||
2240 GATGATGAGATGGCTCTTAAGACCAACCAAGAAATGCTCAATAGCATGACACCTAGCCAGAG 2299
Db |||||
697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
Qy |||||
2300 GTCTGGTACAGAGACTTCATGCTCAGCTCATCAACCAACCCCAATCTCAACAGTGTAG 2359
Db |||||
711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
Qy |||||
2360 TTCTGTGCAACAGTTTGGAAAGGACCGAAGCAACAGCTCGCAAGGCCAGGACATACC 2419
Db |||||
730 TyrCysGluGlnMetTyrPheArgGluLysArgGlnArgGlnLysGlyGly 747
Qy |||||
2420 CCAGGCAACAGTAACAAATGGAAGCCTTCAAGAAATAAGAGAAATAGAGAGGTAGAAACAGGAG 2479
Db |||||
748 -----ProLysTyrLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
Qy |||||
2480 ACCAC---GAATTTGAGAGGCGCCAGGAGTGTCT 2512
Db |||||
764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 9

US-10-175-738-348

; Sequence 348, Application US/10175738

; Publication No. US2003002294A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C45

; CURRENT APPLICATION NUMBER: US/10/175,738

; PRIORITY FILING DATE: 2002-06-19

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348

; LENGTH: 777

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-738-348

Alignment Scores:

Pred. No.:	1,05e-199	Length:	777
Score:	2245.50	Matches:	413
Percent Similarity:	73.67%	Conservative:	141
Best Local Similarity:	54.92%	Mismatches:	177
Query Match:	45,86%	Indels:	21
DB:	15	Gaps:	9
US-09-774-490-1 (1-2709) x US-10-175-738-348 (1-777)			
QY	275	AAGAAACAATGTCCCAAGCTGAAATATCTTACAAAGAAATGTTGGAATCCAAACAATGTG	334
Db	39	LysGlnAsnIleProArgLeuLysLeuThrTyrlsAspLeuLeuSerAsnSerCys	58
QY	335	ATCACTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGAA	394
Db	59	IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu	78
QY	395	CGAGTAGGCTGTATGTTGGACAAGAGTACATATTTTCATTTCGACTGGTTAATATC	454
Db	79	ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu	98
QY	455	---AAGGATTTTCAAGAATCTGTGGCAGTATCTTACACAGAGAGATGAATGCAAG	511
Db	99	AsnLysAsnPheLysLysIleTyrlProAlaAlaLysGluArgValGluLeuCysLys	118
QY	512	TGGGCTGGAAAGACATCTCGAAGATGTGCTAATTTCACTCAAGGTACTTAAAGCATAT	571
Db	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyrl	138
QY	572	AATCAGACTCACTTTCAGCTGTGGAAGCGGGCTTTTCATCCAAATTTGACCTACATT	631
Db	139	AsnLysThrHisIleTyrlValCysGlyThrGlyAlaPheHisProIleCysGlyTyrlle	158
QY	632	GAAATGGACATCATCTCGAGACAAATATTTTAAAGCTGGGAACTCATTTTCAAAAC	691
Db	159	AspLeuGlyValTyrlLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer	178
QY	692	GGCCTGGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGA	751
Db	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu	198
QY	752	GAAATATCTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTCCTATCTTCGAACT	811
Db	199	TyrlLeuTyrlSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer	218
QY	812	CTTGGG-----CACCAACCACTCATCAGACAGACAGCAGTATTCAGGTGGCTC	862
Db	219	LeuGlyProThrHisAspHisIleTyrlleArgThrAspIleSerGluHisTyrlTrpLeu	238
QY	863	AATGATCAAAGTTTCATGTAGTCCCACTCTCATCTCAGAGAGTGACAATCTTGAAGATGAC	922
Db	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrlAsnProAspAspAsp	258
QY	923	AAAGTATATCTTTCTTCGTTGAAATGCAATAGATGAGAACACTCTGGAAAGCTACT	982
Db	259	LysIleTyrlPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle	278
QY	983	CACGCTAGATAGGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAACTCTGTGTAAT	1042
Db	279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn	298
QY	1043	AAATGGACAATCTCTCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGCAAT	1102
Db	299	LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla	318
QY	1103	GACATCATTTTGAAGTCACTGAGATGATTTTCCTTAATGAATTTTAAAGATCTTAAAT	1162
Db	319	AspThrTyrlPheAspGluLeuGlnAspIleTyrlLeuLeuProThrArgAspGluArgAsn	338
QY	1163	CCAGTTGTATGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATCAGCGTGTGT	1222
Db	339	ProValValTyrlGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys	358

QY	1223	ATGTATAGCATGATGATGTGAGAGGTGTCTTGGTCCATATGCCCACAGGATGA	1282
Db	359	ValTyrlSerMetAlaAspIleArgAlaValPheAsnGlyProTyrlAlaHisLysGluSer	378
QY	1283	CCCAACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCACGCCCGAGAACTTGT	1342
Db	379	AlaAspHisArgTrpValGlnTyrlAspGlyArgIleProTyrlProArgProGlyThrCys	398
QY	1343	CCCAAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGACCTTCTCTGATGATGTATA	1399
Db	399	ProSerLysThrTyrlAspProLeuLysSerThrArgAspPheProAspValIle	418
QY	1400	ACCTTTGCAAGAGTCTATCCAGCCATGACATCCAGTTTCTTATGAAACAATCCGCA	1459
Db	419	SerPheIleLysArgHisSerValMetTyrlLysSerValTyrlProValAlaGlyLysPro	438
QY	1460	ATAGTGATCAAAACGGATGTAATATCAATTTTACAAAATTTGCTAGACCGAGTGGAT	1519
Db	439	ThrPheLysArgIleAsnValAspTyrlArgLeuThrGlnIleValValAspHisValIle	458
QY	1520	GCAGAAGATGCACAGTATGATGTTTATCGGAACAGATGTTGGACCGCTTCTTAAA	1579
Db	459	AlaGluAspGlyGlnTyrlAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys	478
QY	1580	GTAGTTTCAATTCCTAAGAGACCTTGTATGATTTAGAGAGGTTCTCTCGAAGAAATG	1639
Db	479	ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluLeu	497
QY	1640	ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA	1699
Db	498	GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln	517
QY	1700	CTATATATTTGTTCAACGGCTGGGTGCCAGCTCCCTTTACACCGGTGTGATTTTAC	1759
Db	518	LeuTyrlleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyrl	537
QY	1760	GGGAAGAGGTGTCTGATGTTGCTGCCCGAGACCTTACTGTCTGGATGGTGTCT	1819
Db	538	GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrlCysAlaTrpAspGlyAsn	557
QY	1820	GCATGTTCTCGTATTTTCCACTGCCAAGAGAGCGACAGACGACAGATATAAGAAT	1879
Db	558	AlaCysSerArgTyrlAlaProThrSerLysArgArgAlaArgGlnAspValLysTyrl	577
QY	1880	GGACACCACTGACTCTACTGTTTCAGACTTACCATGATATCACCATGGCCACAGCCCT	1939
Db	578	GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla	596
QY	1940	GAAGAGAAATCATCTATGTGTGAGAATAGTAGACATTTTGGATGAGTCCGAG	1999
Db	597	AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys	616
QY	2000	TCGACAGAGCGGTGCTTATTCGCAATTCAGAGGCGAAATGAAGCGGAAAGAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTrpTyrlleGlnArgSerGlyAspGluHisArgGluGlu	636
QY	2060	ATCAGGTGATGATCATATCATCAGACAGATCAAGCGCTTCTGTCTAGCTAGTCTCAA	2119
Db	637	LeuLysProAspGluArgIleIleLysThrGluTyrlGlyLeuLeuIleArgSerLeuGln	656
QY	2120	CAGAAGATTCAGCAATACCTCTGCTCGGTGGAACATGGGTTCATACAAACTCTT	2179
Db	657	LysLysAspSerGlyMetTyrlTyrlCysLysAlaGlnGluHisThrPheIleHisThrIle	676
QY	2180	CTTAAGTAACTTGGAGTCAATTCATGACAGACATTTGGAAGAACTTCTTCAATAAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
QY	2240	GATCATGGAGTGGCTCTAAGACCCAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
QY	2300	GTCTGTACAGAGACTTTCATGACGCTCATCAACCCCACTCTCAACACGATGGATGAG	2359


```

QY 1640 ACAGTTTTCGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACCTAAGCAGCAACAA 1699.
Db 498 GlnIlePheLysHisSerIleLeuAenMetGluLeuSerLeuLysGlnGlnGln 517
QY 1700 CTATATATTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
Db 518 LeuTyrlleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTGAGTGTGCTCGCTCGCCGAGACCCCTTACTGTGCTTGGGATGTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyAsn 557
QY 1820 GCATGCTTCGCTATTTCCTCCACTGCAAGAGACGCAACAGCAGCAGATATAAGAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTGCTTTCAGACTTACACCATGATATACCATGGCCGACGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGAATCATCTATGGTGTAGAGATAGTACACATTTTGGAAATGCGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGCGCTGGTCTATTGGCAATTCAGACGCGAATTCAGACGCGAAAGCAAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTyrTrpIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGAGTGGATGATCATATATCATCAGGACAGATCAAGGCTTCTCTAGCTAGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleLysThrGlyLeuLysThrGlyLeuLysThrPheIleHisThrIle 656
QY 2120 CAGAAGGATTCAAGCAATTTACCTCGCATCGGTGGAACTGGGTTCATACAAACTCTT 2179
Db 657 LysLysAspSerGlyWetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTCGAAGTATTGACACAGAGCAATTTGGAAGAACTCTTCATTAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGCGCTCATCAACACCCCAATCTCAACACAGCTGATGACG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TCTGTGAACAGTTTGGAAAGGACCGAAACAACTCGGCAAGCCCGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACAGTAACAAATGGAAGCATTACAGAAATAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCAC-----GAATTGAGAGGCGACCCAGGAGTGTG 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

```

RESULT 11

```

US-10-176-482-348
; Sequence 348, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria

```

```

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIORITY FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-176-482-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-176-482-348 (1-777)
QY 275 AAGAACAATGCGCAAGGCTGAATTTATCTTACAAAGAAATGTTGGAATCCAAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGCTGTGGCCACAGCTCCAGTTATCATATCATCTTCTCTTTCGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuAspGluGlu 78
QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuLys 118
QY 512 TGGGCTGGAAGAAGACATCTCGAAAGATGTCTAATTTTCATCAAGGTACTTAAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCATCTGTAGCTGTGGACCGGGCTTTTCATCCAATTTGCACCTACATTT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGACATCATCTCGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GGCGGTGGGAGAGTCCATATGACCTTAAGCTGTGACGATCCCTTTTAAATAGATGGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGACTTTGCTATCTTCGGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCCCAATCAGGACGAGCAGCAGCATGATTCACAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTCATTAGTCCACCTCATCTCAGAGAGTGAACAATCTCTGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAGTATATCTTTTCTTCTCGTGAATGCAATAGATGAGAGCAACTCTCGGAAAGCTACT 982
Db 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278

```


Db 39 LysGlnAsnIleProArgLeuLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTGGCCACAGCTCCAGTTATCATACCTCTCTTTGGATGAGGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CCGAGTAGCTGTATGTTGGAGCAAGAGATCATATTTTCATTTCGACCTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTCAAAGATTGTGGCCAGCTATCTTACACAGAGAGATGAATGCCAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTGGAAAAGACATCTCGAAAGATGTGTAATTTTCATCAAGTACTTAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTTGTACCCCTGTGGACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTTGGACATCTCTCAGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCCCTGGGAAGAGTCCATATGACCTAAGCTGCTGACAGCATCCCTTTTAAATAGAGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATATATCTCTGGAACCTGACTGATTTTATGGGCGAGACTTTGCTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCAACCCCAATCAGACAGACGACCATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAACTCCTCAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAGATATCTTTTCTTCGTAATAATGCAATAGATGGAGAACACTCTGGAAAAGCTACT 982
Db 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CAGCTAGATAGTGCAGATGCAAGATGACTTTGGAGGCAAGAGTCTGTGTGAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGCAT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACATCACTTTGATGAAGTGCAGATGATATTCCTTAATGAATTTAAAGATCTTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATATGGAGTGTTCAGACTTCCAGTACATTTTCAAGGATCAGCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGATGATGAGAAGGTGTTCCTTGTGTCATATGCCCACAGAGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTCAATGGGTGCTTATCAAGAAAGAGTCCCTATTCACGCCAGGAACTTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCCAGCAACAACTTTGGTGGT---TTTGACTCTCAAGAGGACCTTCTCTGATGATGTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle 418

QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAATCCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGTAGACCGATGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAAGATGCACAGTATGATGTTATCCGACACAGATGTTGGACCGTCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGGACACTTGGTATGATTAATTAGAAAGAGTTCTCTCGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTGTCTATTTCAGCAATGGAGCTTTTCCACTAAGCAGACAA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTGCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GCGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTGCTTGGATGGTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCCTATTTTCCACTGCAAGAGACGACACAGACGACAAAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCACTGACTCCTACTGTTACAGCTTACACCATGATATACCATCGGCCACGACCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAAGAGATCATCTATGTTAGAGAAATAGTAGCACATTTTGGATGAGTCCGGAAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGAAATGAAGACGCAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTCAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTCAGGCAATTAACCTCTGCCATCGCGTGAACATGGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGTAACTTCGGAAGTCAATTCACAGAGCATTTGGAAGAACTTCTTTCAATAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCAGAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTTCATGCTAGCTCATCAACCCACCTCAATCTCAACAGCATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTCAACAGTTTCGAAAAGGACCGAAACACAGCTCGGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACAGTAACAATCGAAGCACTTACAGAAATTAAGAAAGTAGAAGACAGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763

QY 2480 ACCCAC---GAATTTGAGAGGACCCAGGAGTGTC 2512
 ||| : : : : : ||| : : : : :
 Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 13

US-10-176-913-348

; Sequence 348, Application US/10176913

; Publication No. US20030022298A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C66

; CURRENT APPLICATION NUMBER: US/10/176,913

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See file Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348

; LENGTH: 777

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-913-348

Alignment Scores:

Pred. No.: 777
 Score: 1,05e-199 Length: 777
 2245.50 Matches: 413
 Percent Similarity: 73.67% Conservative: 141
 Beat Local Similarity: 54.92% Mismatches: 177
 Query Match: 45.86% Indels: 21
 DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-176-913-348 (1-777)

QY 275 AAGACAATGTCAGGCTGAAATATCTCAAGAAGAAATGTCGAATCCAAATGTG 334
 ||| : : : : : ||| : : : : :
 Db 39 LysGlnAsnIleProArgLeuLysLeuThrLysAspLeuLeuLeuSerAsnSerCys 58
 ||| : : : : : ||| : : : : :
 QY 335 ATCACTTTCATCAATGGCTGGCCACACAGCTCCAGTTATCATACCTCTCTTTGGATGAGGAA 394
 ||| : : : : : ||| : : : : :
 Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
 ||| : : : : : ||| : : : : :
 QY 395 CGGAGTAGCTGTATGTTGGAGAAAGGATCATATTTTCATTCGACCTGTTAATATC 454
 ||| : : : : : ||| : : : : :
 Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 ||| : : : : : ||| : : : : :
 QY 455 ---RAGGATTTCAAGAAGATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
 ||| : : : : : ||| : : : : :
 Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
 ||| : : : : : ||| : : : : :
 QY 512 TGGGCTGGAAAAGACATCTCGAAAGAAATGTGCTTAATTTTCATCAAGTACTTAAGGCATAT 571
 ||| : : : : : ||| : : : : :
 Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
 ||| : : : : : ||| : : : : :
 QY 572 AATCAGACTCATGTTGACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
 ||| : : : : : ||| : : : : :
 Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
 ||| : : : : : ||| : : : : :
 QY 632 GAAATGGACATCATCTCGAGACATATTTTAAGCTCGGAGACTCACATTTTGAAC 691
 ||| : : : : : ||| : : : : :
 Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
 ||| : : : : : ||| : : : : :
 QY 692 GGCCTGGGAAGAGTCCATATGACCCCTAAAGCTGCTGACAGCATCCCTTTTAATAGATGGA 751
 ||| : : : : : ||| : : : : :
 Db 692 GGCCTGGGAAGAGTCCATATGACCCCTAAAGCTGCTGACAGCATCCCTTTTAATAGATGGA 751

Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 ||| : : : : : ||| : : : : :
 QY 752 GAATATATCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGGAAT 811
 ||| : : : : : ||| : : : : :
 Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
 ||| : : : : : ||| : : : : :
 QY 812 CTTGGG-----CACCCACCCCAATCAGGACAGACAGACAGCATGATTCAGAGTGGCTC 862
 ||| : : : : : ||| : : : : :
 Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
 ||| : : : : : ||| : : : : :
 QY 863 AATCATCCAAAGTTCATTAGTGGCCCACTCATCTCAGAGAGTGAACAATCTCGAAGATGAC 922
 ||| : : : : : ||| : : : : :
 Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
 ||| : : : : : ||| : : : : :
 QY 923 AAAGTATATCTTTTCTTCCGTAAGAAATGCATAGATCGAGACACTCTCGGAAAGCTACT 982
 ||| : : : : : ||| : : : : :
 Db 259 LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
 ||| : : : : : ||| : : : : :
 QY 983 CACGCTAGATAGCTCAGATATGCAAGAAATGACATTTTGGAGGGCCACAGAAAGTCTGCTGAAT 1042
 ||| : : : : : ||| : : : : :
 Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
 ||| : : : : : ||| : : : : :
 QY 1043 AAATGGAACAATCTCTCAAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAAATGGCAT 1102
 ||| : : : : : ||| : : : : :
 Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
 ||| : : : : : ||| : : : : :
 QY 1103 GACACTCATTTTGTGATGAACCTCAGAGATGATTCCTAATGAACCTTTAAAGATCCCTAAAT 1162
 ||| : : : : : ||| : : : : :
 Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
 ||| : : : : : ||| : : : : :
 QY 1163 CCAGTTGTATATGAGTGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTGTGT 1222
 ||| : : : : : ||| : : : : :
 Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
 ||| : : : : : ||| : : : : :
 QY 1223 ATGTATAGATGATGATGTCAGAAAGGTGTCTTGTGCTCATATGCCACAGAGATGGA 1282
 ||| : : : : : ||| : : : : :
 Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
 ||| : : : : : ||| : : : : :
 QY 1283 CCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACTGT 1342
 ||| : : : : : ||| : : : : :
 Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
 ||| : : : : : ||| : : : : :
 QY 1343 CCCAGCAAAAACATTTGGTGGT---TTTGACTCTCAAAAGAGACCTTCTCTGATGATGATATA 1399
 ||| : : : : : ||| : : : : :
 Db 399 ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle 418
 ||| : : : : : ||| : : : : :
 QY 1400 ACCTTTGCAAGAAGTCTATCCAGCCATGATCAATCCAGTGTTCCTATGAAACATGCCCA 1459
 ||| : : : : : ||| : : : : :
 Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
 ||| : : : : : ||| : : : : :
 QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
 ||| : : : : : ||| : : : : :
 Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
 ||| : : : : : ||| : : : : :
 QY 1520 GCAGAAGATGCACAGTATGATGTTTATCGGAAACAGATGTTGGACCGTCTCTTAAA 1579
 ||| : : : : : ||| : : : : :
 Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
 ||| : : : : : ||| : : : : :
 QY 1580 CTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCCTCTCGAAGAAATG 1639
 ||| : : : : : ||| : : : : :
 Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
 ||| : : : : : ||| : : : : :
 QY 1640 ACAGTTTTTCCGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACATTAAGCAGCAACAA 1699
 ||| : : : : : ||| : : : : :
 Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
 ||| : : : : : ||| : : : : :
 QY 1700 CTATATATTTGTTCAACCGCTGGGTGCTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
 ||| : : : : : ||| : : : : :
 Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
 ||| : : : : : ||| : : : : :
 QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTGCCCGAGACCTTACTGTGCTTGGATGGTCT 1819
 ||| : : : : : ||| : : : : :
 Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
 ||| : : : : : ||| : : : : :
 Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557

QY 1820 GCATGTTCTCGTATTTTCCCAAGAGACGCAAGAGACGCAAGATTAAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTCACTCACTGTTCCAGACTTACACCATGATATACCATGGCCACACGCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleHis---GluThrAla 596
QY 1940 GAAGAGAGAAATCATCTATGCTAGAGAAATAGTACACATTTTGAATGCAAGTCCGAAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCCGAGAGACGCTGCTATTTGGCAATTCAGAGCGGAAATGAAGAGCGGAAAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGAGTGCATCATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTCAGCAATACCTCTGCCATGCGGTGGAACATGGGTTCATACAAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAGGTAACTGGAAGTCACTGACACAGAGCATTTTGGAAAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAAGTCCATAGCATCACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGAGCTCATCAACACCCCAATCTCAACACGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACAAACGTCGCAAGGCGGAGCATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGAGACAGTAAACAAATGGAAGACATTCACAGAAATTAAGAAAGTAGAAGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTCCAGAGGCGCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 14

US-10-180-552-348
; Sequence 348, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348

; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-180-552-348 (1-777)

QY 275 AAGAACATATGTCACAGGCTCAAAATATCTCTACAAAGAAATGTTGGAATCCAAACATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCATCTTCAATGCTTGGCCACACAGCTCCAGTTATCATACATCTCTCTTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuAspGluGlu 78
QY 395 CGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATTCGACCTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTGGAAGACATCTCTGAAAGATGTCTTAATTTTCATCAAGGTACTTAAGGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTGTAGCCCTGTGGAACGGGGCTTTTCATCCATTTTGCACCTACATTT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCGCTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATCGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTTGTCTATCTTCGAACT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCCCAATCAGGACAGCAGCATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCTGGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAAGTATACCTTTTCTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAAGAGCTACT 982
Db 259 LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CACCTAGATAGTACATATGCAAGATGACATTTTGGGGCGCAGAGTCTCGTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGGACACATTCCTCAAGCTGCTGCTGATTTTGTCTCAGTGCCAGGTGCCAAATGGCATT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTGATGAACCTGCAGGATGTATTCCTAATGAACCTTTAAAGATCTCTAAAT 1162
Db 1103 GACACTCATTTGATGAACCTGCAGGATGTATTCCTAATGAACCTTTAAAGATCTCTAAAT 1162

Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CAGATTGATATGAGAGTGTTCAGACTCCAGTAACATTTTCAAGGATCAGCCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLeuGlySerAlaValCys 358
QY 1223 ATGTATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCACAGGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysSer 378
QY 1283 CCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAATGTT 1342
Db 379 AlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCACAGAAACATTTGGTGGT---TTTGACTCTACAAGGACCTTCCTGATGATGTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCGCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTATCAAAACCGGATTAATTAATTAACAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAGATGGACAGTATGATTTATGTTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAAAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTyrPheAsnMetGluValValLeuGluLeu 497
QY 1640 ACAGTTTTCGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 498 GlnIlePheLysHisSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTTGTTCAACGGCTGGGTGCGGCTCCAGTCCCTTTACACGGTGTGATTTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGCTTGGAGTGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysLysLeuAlaArgAspProTyrCysAlaThrAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATAAGAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTACTCAGTGTTCAGACTTACACCATGATATACCATGCGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATAGTAGCACATTTTGAATGCACTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGCGCTGCTGTATTGGCAATTCAGAGCGGAAATGAAGAGCAAAAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTyrPheGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGTCTACNA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTTCAGGCAATTCCTGCTCCATGCGTGGAAATCGGTTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAAGTCAATTCACACAGACATTTTGAAGAACTTCTTCATAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696

QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCTCAATAGATGACACCTAGCCAGAAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGTACAGAGACTTTCATCAGCTCATCAACACCCCAATCTCAACACGATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGACCGAAACCAACGTCGCGCAAGCCGAGACATACC 2419
Db 730 TyrCysGluGlnMetTyrHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGACACAGTAAACAATGGAAGCAGCTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTyrIleHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGCGCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
RESULT 15
US-10-180-557-348
; Sequence 348, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180.557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-348
Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
Gaps: 9
US-09-774-490-1 (1-2709) x US-10-180-557-348 (1-777)
QY 275 AAGAACAAATGTGCAAGGCTGAAATTTATCTTACAAAGAAATCTTGAATCCAAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGCTTGGCCCAACAGCTCCAGTTCATACCTCTCTTTGGATGAGNA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuAspGluGlu 78
QY 395 CGGAGTAGCTGTATGTTGGAGCAAGGATCACAATATTTTCATTCGACCTGTTAATATC 454
Db 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGAGATGAATGCAAG 511

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:26 ; Search time 417 Seconds
(without alignments)
11308.390 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 4896
Sequence: 1 aatctttttttatcatgatg.....agggtttttttcttaataacc 2709

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 11160482

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-DB=US09774490/runat_03082003_095417_25381/app query fasta_1.2887
-DB=Pending_Patents_AA_Main -QFMT=fastan -SUFFIX=rapm -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi
-LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1.628 @runat_03082003_095417_25381 -NCPU=6 -ICPU=3
-NO_MMAP -LARGUEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents AA Main:
1: /cgn2_6/ptodata/2/paa/PCTUS COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06 COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07 COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08 COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US081 COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US082 COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US083 COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US084 COMB.pcp.*
9: /cgn2_6/ptodata/2/paa/US085 COMB.pcp.*
10: /cgn2_6/ptodata/2/paa/US086 COMB.pcp.*
11: /cgn2_6/ptodata/2/paa/US087 COMB.pcp.*
12: /cgn2_6/ptodata/2/paa/US088 COMB.pcp.*
13: /cgn2_6/ptodata/2/paa/US089 COMB.pcp.*
14: /cgn2_6/ptodata/2/paa/US090 COMB.pcp.*
15: /cgn2_6/ptodata/2/paa/US091 COMB.pcp.*
16: /cgn2_6/ptodata/2/paa/US092 COMB.pcp.*
17: /cgn2_6/ptodata/2/paa/US093 COMB.pcp.*
18: /cgn2_6/ptodata/2/paa/US094 COMB.pcp.*
19: /cgn2_6/ptodata/2/paa/US095 COMB.pcp.*
20: /cgn2_6/ptodata/2/paa/US096 COMB.pcp.*
21: /cgn2_6/ptodata/2/paa/US097A COMB.pcp.*
22: /cgn2_6/ptodata/2/paa/US097B COMB.pcp.*
23: /cgn2_6/ptodata/2/paa/US098 COMB.pcp.*
24: /cgn2_6/ptodata/2/paa/US099A COMB.pcp.*
25: /cgn2_6/ptodata/2/paa/US099B COMB.pcp.*
26: /cgn2_6/ptodata/2/paa/US100 COMB.pcp.*
27: /cgn2_6/ptodata/2/paa/US101 COMB.pcp.*
28: /cgn2_6/ptodata/2/paa/US102 COMB.pcp.*

29: /cgn2_6/ptodata/2/paa/US103 COMB.pcp.*
30: /cgn2_6/ptodata/2/paa/US104 COMB.pcp.*
31: /cgn2_6/ptodata/2/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4201	85.8	771	1	PCT-US00-41943-24	Sequence 24, Appl
2	4201	85.8	771	1	PCT-US02-07826-284	Sequence 284, App
3	4201	85.8	771	8	US-08-488-935-2	Sequence 2, Appli
4	4201	85.8	771	8	US-08-489-057-2	Sequence 2, Appli
5	4201	85.8	771	20	US-09-694-085-24	Sequence 24, Appl
6	4201	85.8	771	22	US-09-791-537-31106	Sequence 31106, A
7	4201	85.8	771	26	US-10-067-632-54	Sequence 54, Appl
8	4201	85.8	771	26	US-10-097-340-284	Sequence 284, App
9	4201	85.8	771	28	US-10-247-671-164	Sequence 164, App
10	4201	85.8	771	28	US-10-362-538-10	Sequence 10, Appl
11	4201	85.8	771	31	US-09-791-537-101922	Sequence 101922, A
12	4057.5	82.9	772	22	US-09-791-537-72548	Sequence 72548, A
13	4035.5	82.4	772	22	US-09-362-167-32	Sequence 32, Appl
14	3989.5	81.5	772	22	US-09-791-537-11919	Sequence 11919, A
15	3989.5	81.5	772	22	US-09-791-537-128280	Sequence 128280, A
16	3784.5	77.3	772	22	US-09-791-537-62757	Sequence 62757, A
17	3478.5	71.0	666	22	US-09-791-537-105918	Sequence 105918, A
18	3207.5	65.5	777	19	US-09-582-419A-1	Sequence 1, Appli
19	2245.5	45.9	777	24	US-10-006-041A-310	Sequence 310, App
20	2245.5	45.9	777	24	US-10-006-063A-310	Sequence 310, App
21	2245.5	45.9	777	26	US-10-006-116A-310	Sequence 310, App
22	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
23	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
24	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
25	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
26	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
27	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
28	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
29	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
30	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
31	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
32	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
33	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
34	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
35	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
36	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
37	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
38	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
39	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
40	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
41	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
42	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
43	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
44	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
45	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App

ALIGNMENTS

RESULT 1
PCT-US00-41943-24
; Sequence 24, Application PC/TUS0041943
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441W01
; CURRENT APPLICATION NUMBER: PCT/US00/41943
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/694,085
; PRIOR FILING DATE: 2000-10-20

! PRIOR APPLICATION NUMBER: US 60/164,056
! PRIOR FILING DATE: 1999-11-08
! NUMBER OF SEQ ID NOS: 48
! SOFTWARE: FASTSEQ for Windows Version 4.0
! SEQ ID NO 24
! LENGTH: 771
! TYPE: PRT
! ORGANISM: Homo sapiens
PCT-US00-41943-24

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-24 (1-771)

QY 200 ATGGGCTGTTAACTAGGATTTCTGTCTTTCTGGGAGTATTACTTACACAGAGCA 259
DB 1 MetGlyTrpLeuThrArgileValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGCTGMAATTATCTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTTCGATCAGAACCGGATAGGCTGTATGTGGGCAAGAGATACATATTTTCATTC 439
DB 61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAATGGGCTGGAAGACATCTCTGAAAGATGTGCTAAATTTTCATCAAGGTA 559
DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCCTGTACGCTGTGGACCGGGCTTTTCATCCAATT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAAATTCGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAACGGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTACTCTGGAAGTGGAGTATTTTATGGGCGAGACTTTGCT 799
DB 181 LeuIleAspGlyLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACCTCTTGGGCACACACCAATCAGACAGACAGCATGATTCCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTTCAATGATCCAAAGTTTATAGTCCCACTCATCTCAGAGAGTGACAATCCTCAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGTGAAATGCAATAGATGGAGACACTCTGGAAGCT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTTGGAGGGGCACAGAAAGTCTGTTG 1039

DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACAATCTCTCAAAAGCTCTGTGATTGCTCAGTCCAGGTCCTCAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTGTTGATGAATCAGGATGTATTCTTAATGAATCTTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGGAGTGTATTCAGCACTTCCAGTAACATTTTCAAGGATACCCGTG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATATCAGCATGATGTGAGAAGGTGTCTTGTGCTCATATGCCACAGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTCACTCTCAAGAGACCTTCTCTGATGATGTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAACTCCAGTGTTCCTATGAAACAATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTACAAAATTTGCTGAGACCGAGTGGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGCAAGTATGATGTTATGTTTATCGGAACAGATGTGGGACCGTCTTTAAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGACACTTGTATGATTATAGAGAGTCTCTCTGGAAGAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGGAACCGACTGTATTTTACCAATGAGACTTTCACCTAAGCAGCAACAA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGGTGTGCTGAGTGTGGCTCGCCGAGACCTTACTGCTGTGGATGGTGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACCAAGATATAGAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTTACAGCTTACCATGATATACCATGGCCACAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGATCATCTATGTTGTAGATAGTACCATTTTGGATGAGTCCGAG 1999
DB 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGGTCTATTGGCAATTCAGAGGCGAAATGAAGACGCAAGAAAGACAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgGlyGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAGTCTCAA 2119

Db 621 lleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy 2120 CAGAGGATTACCAATACCTCTGCCATCGGTGGAAACATGGGTTCATACAAACTCTT 2179
Db 641 GlnlysAspSerGlyAsnTyrLeuGlyHisAlaValGluHisGlyPheIleGlnThrLeu 660
Qy 2180 CTTAGGTAACCTCGAGTCAATCACAGACGAGCAATTTGGAAGACTTCTTCATAAGAT 2239
Db 661 LeuIysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisIysAsp 680
Qy 2240 CATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATCACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerIysThrIysGluMetSerAsnSerMetThrProSerGlnlys 700
Qy 2300 GTCTGGTACAGACTTCATGAGTCAATCAACACCCCAATCTCAACACGATGATGAG 2359
Db 701 ValTprTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
Qy 2360 TTCTGTGACAGATTGGAAAGGACCGAACAACACCTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpIysArgAspArgIysGlnArgGlnArgProGlyHisThr 740
Qy 2420 CCAGGGAACAGTAACAAATGGAAGCACCTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnIysTrpIysHisIleuGlnAsnIysLysGlyArgAsnArgArg 760
Qy 2480 ACCACGAAATTGAGAGGCCACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

PCT-US02-07826-284
; Sequence 284, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07826-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US02-07826-284 (1-771)

Qy 200 ATGGCTGGTTAACTAGGATTGTCGTCTTTCTGGGGAGTATTACTTACAGCAAGAGCA 259

Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
Qy 260 AACTATCAGAAATGCAAGAACAAATGTGCAAGGCTGAAATATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrGlnIleGlnGlyLysAsnValProArgLeuIysLeuSerTyrLysGluMetLeu 40
Qy 320 GAATCCAAATGTGATCATTCTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
Qy 380 CTTTGGATGAGGACCGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
Qy 440 GACCTGGTTAATATCAAGGATTTCAAAGATTGTGGCCAGTATCTTACACAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
Qy 500 GATCAATGCAAGTGGCTGGAAAGACATCTCGAAGAATGTGCTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysIleTyrAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
Qy 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
Qy 620 TGCACCTACATTTGAAATTTGGACATCATCTCGAGACAAATATTTTAAAGCTCGAAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
Qy 680 CATTTGAAAACGGCCGTGGGAAGAGTCCATATGACCTTAACTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
Qy 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACACTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCCGAATCTTGGGCACACCCCAATCAGGACAGACAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisIleProIleArgThrGluGlnHisAspSerArgTrp 220
Qy 860 CTCAATGATCAAAAGTTTCATTAGTCCCACCTCATCTCAGAGAGTGACAAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
Qy 920 GACAAAGTATACTTTTCTCCGTGAAAATGCAATAGATGAGAACACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
Qy 980 ACTCACGCTAGATAGTACAGATATGCAAGATGACTTTGGAGGGCACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
Qy 1040 AATAAATGGACAACATTTCTCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATAGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
Qy 1100 ATTGACACTCATTTTGTAGAACTGACAGATGTATCTTAATGAACCTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy 1160 AATCCAGTCTGTATATGAGTGTTCACGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
Qy 1220 TGTATGTATAGCATGAGTGTATGTGAGAGGGTGTCTCTGCTCCATATGCCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
Qy 1280 GGACCCCAACTATCAATGGGTGCTTATCAAGGAGAGTCCCTTATCCACGCCCAAGAACT 1339

Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAACAACTTTGGTGGTTTGGACTCTACAAAGGACCTTCCTGATGATCTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACCGATTAATATCAATTTACAAATTTCTGTAGACCGATGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAGATGCGACATGATCTATGTTTATCGGACAGATGTTGGACCGTCTCTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGAACCGACTCTATTTCCAGAAATGGAGCTTTCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTTGGTCAACGCTGGGTTGCCAGCTCCCTTTACACGGGTGATATTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTGGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTCTTCGTATTTTCCCACTGCAAGAGACCCAGACGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 CGAGACCCACTCACTTCCTGCTTACACCTTACACCATGATATACCATGGCCACAGCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATCGTGTAGAGATAGTACACATTTTTCGATCGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGAGAGACGCTGTCTATTGGCAATTCAGAGCGCAATGAAAGAGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGATGCGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATTAACCTTCGCCATGCGGTGGAACATGGGTTCATACAACTCT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCGAAGTCAATTGACAGAGCAATTCGAAAGAACTTCTTCATAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGTGGCTCTTAAGCCCAAGAAATGTCNAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGTACAGAGATTCATGACAGCTCATCAACCCCACTCTCAACACAGATGATGAG 2359
Db 701 ValTyrTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGGACCCGAAACACAGCTCGCGCAAGGCCAGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

QY 2420 CCAGGCAACAGTAACAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACCAATTTGAGAGGCGCACCCAGGAGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3

US-08-488-935-2
; Sequence 2, Application US/08488935
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/488,935
; APPLICATION NUMBER: US/08/488,935
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/489,057
; FILING DATE: 09-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-935-2

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 8 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-488-935-2 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATGTCTGTTTCTGGGGAGTATTCTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCATGATCGGACAGCAACAAATGTCCAAGGCTGAATATCTCTACAAAGAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

320 GAATCCAACTGATGATCTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
Db |||||SerAsnVal|||ThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATTC 439
Db |||||LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
440 GACCTGGTTAAATATCAAGGATTTCAAAGATTTGTGTGCCAGTATCTTTACACCAAGA 499
Db |||||AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArg 100
500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAGAATGCTAAATTTTCATCAAGTA 559
Db |||||AspGluCysLysTyrAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
560 CTTAAGGCATATAATCAGACTCACCTGTACGCTGTGGAAACGGGGCTTTTCATCCAAT 619
Db |||||LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
620 TGCACCTACATTTGAATGGACATCATCTCTGAGCAATATTTTAAAGCTGGAGAATCA 679
Db |||||CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
680 CATTTTGAAGCGGCTGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
Db |||||HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTGGAAGTCTGAGCTGATTTTATGGGGAGACTTTCCT 799
Db |||||LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAAGCTTTGGGCACCAACCAATCAGGACAGACAGCATGATTCAGAGTGG 859
Db |||||IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
860 CTCATGATCCAAAGTTCTATAGTCCACCTCATCTCAGAGAGTGACAACTCTGAAGAT 919
Db |||||LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProLysVal 240
920 GACAAAGTATATCTTTCTTCGTCGAAATGCAATGATGAGAGACACTCTGGAAAGCT 979
Db |||||AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGATAGGTGAGATGCAAGATGACTTTGGAGGCGCAGAGTCTGCTG 1039
Db |||||ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAATGGAACAACTTCTCAAAGCTGCTGATTTGCTCAGTCCAGGTCCTCAATGGC 1099
Db |||||AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTGTGAACTGCTGAGGATGATTCCTTAATCAACTTTTAAGATCTTAA 1159
Db |||||IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 AATCCAGTTGTATGAGAGTGTTCAGGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
Db |||||AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGATGTATAGCATGAGTGTGAGAGGTTCTTGTGTCATATGCCACAGGAT 1279
Db |||||CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCACGCCAGGACT 1339
Db |||||GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACTTCTCTGATGATGTTATA 1399
Db |||||CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATGCCCA 1459

401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
1460 ATAGTGATCAAAACGGATGTAAATTTATCAATTTTACAAAAATTTGCTAGACCGAGTGGAT 1519
Db |||||IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAAGATGCACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTTTAAA 1579
Db |||||AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGACACTTGTATGATTTAGAGAGTTCCTCTGGAAGAAATG 1639
Db |||||ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
1640 ACAGTTTTTTCGGGAACCGACTGCTTATTTTCAGCAATGAGGCTTTCACATAAGCAGCAACA 1699
Db |||||ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db |||||LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAAGCTGTCTGAGTGTGCTCGCCCGAGACCTTACTGTCTTGGATGGATGCTTCT 1819
Db |||||GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCAACAGCAGCAAGATATAGAAT 1879
Db |||||AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgLysPheAspIleArgAsn 560
1880 GGAGACCCACTGACTCACTGTTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db |||||GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
1940 GAAAGAGATCATCTATGTTGTTAGAAATAGTAGACATTTTGGAAATGAGTCCGAG 1999
Db |||||GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGAAATGAAGACGCAAAAGAAAG 2059
Db |||||SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgGlyGluGlu 620
2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAA 2119
Db |||||IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
2120 CAGAAGATTTCAGCAATTTACTCTGCGATGCGGTGGAAATGCGTTCATACAACTCTT 2179
Db |||||GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
2180 CTTAAGGTAAACCTGGAGTTCATTGACACAGCATTTGGAGAACTTCTTTCATAAAGAT 2239
Db |||||LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
2240 GATGATGGAGATGCTCTTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db |||||AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
2300 GTCTGGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGAG 2359
Db |||||ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
2360 TTCTGTCAACAAGTTTCGAAAAGGCGGACCAAAACAACTCGGCAAGCCAGGACATACC 2419
Db |||||PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
2420 CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAATAAGAAAGGTAGAAAACAGAGG 2479
Db |||||ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
2480 ACCCAAGAAATTTGAGAGGCGACCCAGGAGTGTCT 2512

```

Db      761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
US-08-489-057-2
; Sequence 2, Application US/08489057
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSES: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/489,057
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-489-057-2
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 8 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-489-057-2 (1-771)
QY      200 ATGGGCTGTTAACTAGGATTCCTCTTTCTGGGAGTATTACTACAGCAAGACA 259
Db      1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY      260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATATTCCTACAAAGAAATGTTG 319
Db      21 AsnTyrGlnAsnGlyLysAsnAsnValProAArgLeuLysLeuSerTyrLysGluMetLeu 40
QY      320 GAATCCAAACATGTGATCACTTTCAATGGCTTGGCCAAAGCTCAGATTATCATCTTC 379
Db      41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY      380 CTTTGGATGAGAGCGGAGTAGGCTGTATGTTGGAGCAAGCATACATATTTTCATTC 439
Db      61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY      440 GACCTGGTTTAAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTTACACCAAGA 499

```

```

Db      81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY      500 GATCAATCCAACTGGCTGGAAAGACATCTGAAAGATGTGCTAAATTTTCATCAAGTA 559
Db      101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY      560 CTTAAGGCATATAATCAGACTCAGCTTGTACCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db      121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY      620 TGCACCTACATTGAAATTTGGACATCCTCAGAGCAATATTTTAACTGGAGAACTCA 679
Db      141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY      680 CATTTTGAAACGGCGCTGGGAAGAGTCCATATCAGCTTACCTTACAGCTGCTCAGAGCATCCCTT 739
Db      161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY      740 TTAATAGATGGAGAAATTATCTCTGGAACCTCAGCTGATTTTATGGGCGGAGACTTTGCT 799
Db      181 LeuIleAspGlyLeuLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY      800 ATCTTCCGAACCTCTTGGGACACACACCAATCAGGACAGAGCATGATTCAGGTGG 859
Db      201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY      860 CTCATGATCCAAAGTTTCATTAGTCCACCTCATCTCAGAGAGTGACAATCTCTCAAGAT 919
Db      221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY      920 GACAAAGTATATCTTTCTCCGTGAAATCAGATAGATGGAGAACACTCTCGAAAAGCT 979
Db      241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY      980 ACTACGCTAGAAATAGGTGAGATATGCAAGAAATGACATTTGGAGGCGCAGAGTCTGTG 1039
Db      261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY      1040 AATAANTGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATGCG 1099
Db      281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY      1100 ATTGACACTCATTTTGTGATGAATCGAGGATGTATTCTTAATCAACTTTTAAAGATCCTAAA 1159
Db      301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY      1160 AATCCAGTTGTATATGGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db      321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY      1220 TGTATGTATAGCATGATGATGAGAGGCTGTCTTGTGTCATATGCCACACAGGAT 1279
Db      341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY      1280 GGACCCCACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCAGGAACT 1339
Db      361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY      1340 TGTCCCGAGAAAACATTTGGTGGTTTGTGACTCTCAAAAGGACCTTCTCTGATGATGTTATA 1399
Db      381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY      1400 ACCTTTGCAGAACGTCATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAACTCGCCA 1459
Db      401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetCAsnAsnArgPro 420
QY      1460 ATAGTGATCAAAACCGATGTAATTTATCAATTTTACAAAATTTGTCTAGACCGAGTGGAT 1519
Db      421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY      1520 GCAGAGATGACAGTATGATGTTTATCGGAACAGATGTTGGACCGCTTCTTAAA 1579
Db      441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

```

```
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
|
|
|
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
|
|
|
QY 1640 ACAGTTTTTCGGGAACCGACTCTATTTCAGCAATGGAGCTTCCACATAGCAGCAACA 1699
|
|
|
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
|
|
|
QY 1700 CTATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGTATATTTAC 1759
|
|
|
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
|
|
|
QY 1760 GGAAGAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTGGATGGTCT 1819
|
|
|
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
|
|
|
QY 1820 GCATGTTCTTCGCTATTTCCTCCACTGCAAGAGACCCACAAAGACGACACAGATATAAGAA 1879
|
|
|
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
|
|
|
QY 1880 GGAGACCCACTCACTCACTGTTTCAGACTTACACCATGATAATCACCATGGCCACACCCCT 1939
|
|
|
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisIleGlyHisSerPro 580
|
|
|
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATAGTAGACACATTTTGGATGCGATGCCGAAG 1999
|
|
|
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
|
|
|
QY 2000 TCGCAGAGAGCGCTGCTATTGGCAATTCAGAGCGCAATGAGAGCGGCAAGAGAGAG 2059
|
|
|
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgGlnGluGluArgLysGlu 620
|
|
|
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACA 2119
|
|
|
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
|
|
|
QY 2120 CAGAAGGATTGAGCAATTCCTCGCATGCGGTGGAAACATGGGTTTCATACAACTCTT 2179
|
|
|
Db 641 GlnLysAspSerGlyAsnTyrLeuGlyHisAlaValGluHisGlyPheIleGlnThrLeu 660
|
|
|
QY 2180 CTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCAATTTGGAGACTTCTTCATTAAGAT 2239
|
|
|
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
|
|
|
QY 2240 GATGATGAGATGGCTCTAAGACCAACCAAAATGTCCAATAGCATGACACCTAGCCAGAAG 2299
|
|
|
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
|
|
|
QY 2300 GTCTGGTACAGAGACTTCATGCGCTCATCAACCCCAATCTCAACACGATGGATGAG 2359
|
|
|
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
|
|
|
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACACCTCGCAAGAGCCAGGACATACC 2419
|
|
|
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
|
|
|
QY 2420 CCAGGGAACAGTAACAAATGGAAGCACTTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
|
|
|
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
|
|
|
QY 2480 ACCCAAGATTGAGAGGCGACCCAGGAGTGTC 2512
|
|
|
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
```

RESULT 5

```
US-09-694-085-24
; Sequence 24, Application US/09694085
; GENERAL INFORMATION:
; APPLICANT: Behar, Oded
; APPLICANT: Woolf, Clifford J.
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441001
```

```
; CURRENT APPLICATION NUMBER: US/09/694,085
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-694-085-24

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.80% Indels: 0
DB: 20 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-09-694-085-24 (1-771)

```
QY 200 ATGGCGTGGTTAACTAGGATTGTCTCTTTCTGGGGAGTATTACTTACAGCAAGACA 259
|
|
|
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
|
|
|
QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTACAAAGAAATGTTG 319
|
|
|
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
|
|
|
QY 320 GAATCCAAATATGTATGATCATTTCATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
|
|
|
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
|
|
|
QY 380 CTTTGGATGAGAACCGAGTAGGCTGTCTGTGGCAAGAGATCATATATTTTCATTC 439
|
|
|
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
|
|
|
QY 440 GACCTGGTTAATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAGAGA 499
|
|
|
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
|
|
|
QY 500 GATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA 559
|
|
|
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
|
|
|
QY 560 CTTAAGCATATATCAGACTCTGTCTGACCTGTGGACGGGGCTTTTCATCCCAAT 619
|
|
|
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
|
|
|
QY 620 TGCACTACATTTGAAATTTGGACATCATCTCGAGACAATATTTTAAAGCTGGAGAACTCA 679
|
|
|
Db 141 CysThrTyrIleGluIleGlyHisIleProGluAspAsnIlePheLysLeuGluAsnSer 160
|
|
|
QY 680 CATTTTGAAACCGCCCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
|
|
|
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
|
|
|
QY 740 TTATATAGATGAGAAATATATCTCTGGAACTGCAGCTGATTTTATGGGCGAGACTTTGCT 799
|
|
|
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
|
|
|
QY 800 ATCTTCCGAATCTTTGGGCAACCCACCCCAATCAGGACAGAGCATGATTTCCAGGTGG 859
|
|
|
Db 201 IlePheArgThrLeuGlyHisHisIleProLysThrGluGlnHisAspSerArgTrp 220
|
|
|
QY 860 CTCATATCATCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
|
|
|
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
|
|
|
QY 920 GACAAAGTATATCTTTTCTTCCTGGTAAATGCAATAGATGAGAGACACTCTGGAAAAGCT 979
|
|
|
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
|
|
|
```

QY 980 ACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGGAGGCGCAGAGTCTGCTG 1039
DB |||||||
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACAACTTCCTCAAGCTCTGCTGATTTGCTCAGTGCAGGTCCTCAATGCG 1099
DB |||||||
DB 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATGACACTCATTTTGTAGAACTGCAGGATGTATTCCTTAATGAATTTAAAGATCCTAAA 1159
DB |||||||
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTATACGACTTCAGCTTACATTTTCAAGGGATCAGCCGTG 1219
DB |||||||
DB 321 AsnProValValTyGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTTCCTTTGGTCCATATGCCACAGGAT 1279
DB |||||||
DB 341 CysMetTySerMetSerAspValArgValPheLeuGlyProTyAlaHisArgAsp 360
QY 1280 GACCCAACTATCAATGGTGCTTATCAAGAGAGTCCCTATCCACGGCCAGAACT 1339
DB |||||||
DB 361 GlyProAsnTyxGlnTrpValProTyxGlnGlyArgValProTyxProArgProGlyThr 380
QY 1340 TGTCCCAAGCAACATTTTGGTGGTGTGACTCTACAAAGGACCTTCCTGTATGATGTTATA 1399
DB |||||||
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAAATCGCCA 1459
DB |||||||
DB 401 ThrPheAlaArgSerHisProAlaMetTyxAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGATTAATATCAATTTACAAATTTGTCGTAGACCGAGTGTAT 1519
DB |||||||
DB 421 IleValIleLysThrAspValAsnTyxGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAGATGGACATGATGATGTTATGTTATCGCAAGAGTGTGGGACCGTCTCTAAA 1579
DB |||||||
DB 441 AlaGluAspGlyGlnTyxAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 TPAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
DB |||||||
DB 461 ValValSerIleProLysGluThrTyxAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTTCCATAGCAGCAACAA 1699
DB |||||||
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTTGTTCAACGGCTGGGTGGTCCAGCTCCCTTTACACGGGTGTATATTAC 1759
DB |||||||
DB 501 LeuTyxIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyx 520
QY 1760 GCGAAGAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCCCTTACTGTGCTTGGGATGTTCT 1819
DB |||||||
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyxCysAlaTrpAspGlySer 540
QY 1820 GATGTTCTCGCTATTTTCCCTGCTGCAAGAGACGCAAGACGACAGATATAAGAAAT 1879
DB |||||||
DB 541 AlaCysSerArgTyxPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCAGTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
DB |||||||
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTAGCAATTTTGGATTCAGTCCGAG 1999
DB |||||||
DB 581 GluGluArgIleIleTyxGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TGCAGAGAGCCGCTGCTATTGGCAATTCAGAGCGGCAATTAAGAGCGGAAAGAGAG 2059
DB |||||||
DB 601 SerGlnArgAlaLeuValTyxTrpGlnPheGlnArgArgAsnGluArgLysGluGlu 620

QY 2060 ATCAGATGGATGATCATATCATCAGACAGATCAAGCGCTTCTGCTAGTAGTACAA 2119
DB |||||||
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGATTACGCAATACCTCTGCATCGGTGGAAACATGGGTTTCAATCAAACTCTT 2179
DB |||||||
DB 641 GlnLysAspSerGlyAsnTyxLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCCGGAGTCAATTGACACAGAGCATTTGGAGAACTTCTTCATAAAGAT 2239
DB |||||||
DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATCATGAGATGCTCTTAAGACCAAAATGTCCTCAATAGCATCAGACACCTAGCCAGAG 2299
DB |||||||
DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGCAGCTCATCAACCAACCCCAATCTCAACACATGGATGAG 2359
DB |||||||
DB 701 ValTrpTyxArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGACCGAAACAACGTCGCGCAAGGCCAGGACATACC 2419
DB |||||||
DB 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGAAGACACTTACAAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
DB |||||||
DB 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCGCCAGGAGTGC 2512
DB |||||||
DB 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 6
US-09-791-537-31106
; Sequence 31106, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 31106
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-31106

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-31106 (1-771)

QY 200 ATGGCGTGTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTACAGCAGAGCA 259
DB |||||||
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATCGGAAGAACAAATGTCGCAAGGTGAAATTTATCTCAAAAGAAATGTTG 319
DB |||||||
DB 21 AsnTyxGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyxLysGluMetLeu 40
QY 320 GAATCCAAATATGATCATTTCATCGCTGGTGGCCCAACAGCTCCAGTTATCATCCTTC 379
DB |||||||
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyxHisThrPhe 60

QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGATCATATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCGTGGAAAAGACATCTCTGAAAGAATGTGCTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysIleTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACCTGTGACGCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATGGAATTGGACATCATCTCTGAGGCAATATTTTAAAGTGGAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATTTACTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTCCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTGGGCGACACCCACCATCCATCAGAGAGTCAATCTCCTCAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTTCAGATGTCAGAAATGATCTTGGAGGCCACAGAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGAACAATCTCTCAAGCTCGTCTGATTGCTCAGTCCAGAGTCCCAATGSC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCAATTTGATGAACTGCAGGATGTATTCCTAATGAACTTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTGTATATGGAGTGTTCAGACTCCAGTAACTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGAGAAGGTGTTCTTGGTCCATATGCCCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTATCAATGGGTGCTTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTTCATCTACAAAGAGACTTCTCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCATGACATCCAGTGTTCCTATCAACAATGCCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420

RESULT 7

QY 1460 ATAGTGATCAAAACGGATGTAAATATCAATTTACAAATTTGCTGTAGACCGATGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGCACAGTATGATGTTATCGGAACAGATGTGGGACCGTCTCTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 CTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCTCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGAAACCGACTGCTATTTACAGCAATGGAGCTTTCACATAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTTCAACGGCTGGGTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAAGCTGCTGAGTGTGCTCGCCCGACACCTTACTGTGTGCTGGATGGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAGCAAGATATAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTCTGTTACAGCTTACACCATGATATCACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAAGAGAATCATCTATGTTAGAGAAATAGTAGACATTTTGGAAATGAGCTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGTGCTATTTGGCAATTCAGAGGCGAAATGAAGCGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACTCTGCTCATCGGTGGAACATGGTTCATACAAACTCT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTTGGAGTCAATTGACACAGAGCATTTGGAAGAACTTCTTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATCATGGAGTGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATGACGCTCATCAACCCACCCCAATCTCAACACCATGATCAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTCGAAAAGGACCGAAAACAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACTTGAAGACCTTACAGAAATTAAGAAAGGTAGNAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCACAATTTGAGAGGGCACCCAGGAGTGTCT 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

US-10-067-632-54

Sequence 54, Application US/10067632
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
Kolodkin, Alex L.
Matthes, David
Bentley, David R.
O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/067,632
FILING DATE: 04-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/835,268
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-067-632-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-067-632-54 (1-771)

QY 200 ATGGCTGTTAACTAGGATTTCTGCTTTCTGGGAGTATTAATTAACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgileValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAAGGAGAACATGTGCCAAGCTGAAATTAATCTACAAAGAAATGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGAACCGGATGAGCTGTATGTGGAGCAAGGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTTGTGTGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACCCCTGTGGAAACGGGGCTTTTCATCCTCAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTCGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAAGCTCTGACAGCATCCCT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTAATCTCTGGAACCTCAGCTGATTTTATGGGCGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGGCACCCACCATCAAGGACAGGACAGCAGTATTCACAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTCCCACTCATCTCAGAGAGTGACAACTCTGGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTTCTCCGTGAAAATCCATATAGATGAGAGACATCTCTGGAAGACT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATATCAAGAAATCACTTTTGGAGGCGACAGAAAGTCTGCTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAAGCTCGTCTGATTTTCTCAGTGCCCAAGTCCAAATGCG 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTATGAATCTCAGAGTGTATTCCTAACTGAACCTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGTGAGAAAGGTGTCTTGTGTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCACGCCCGAGACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGAAAACATTTGGTGGTTTGTACTCTCAAGAGGACCTTCTCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAGAAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTACAAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

```
QY 1520 GCAGAGATGGACAGTATGATGTTATGTTATTCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnPyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGCAGACTGCTATGATTTTGAAGAGAGCTTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpThrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGCTGGGTTGCCAGCTCCCTTTACACCGGTGATATTTAC 1759
Db 501 LeuTyrlleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleYr 520
QY 1760 GCGAAAGCGTGTCTGAGTGTTCCTCGCCGCGAGACCTTACTGTGCTGGGATGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrcysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCGACAGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrcysProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCAGTCTGCTGACATTCACCATGATGATATCACCAGCCGCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCGATCGGAG 1999
Db 581 GluGluArgIleIleTyrlleGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCCGAGAGACCGCTGCTTATTTGGCAATTCAGAGCGGCAAAATCAAGAGCGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrlleGlyValGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTTCAGGCAATTAACCTCTGCCATGGGTGGAAACATGGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrlleGlyValGlnCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACCTGGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCATAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGTGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGCTACAGAGCTTCATGCTCAGCTCATCAACCCCACTCAACAGCATGATGATGAG 2359
Db 701 ValTrpTyrlleArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCGAAAGCGCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAGCACTTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArg 760
QY 2480 ACCCAGCAATTTGAGAGGGCCACCCAGGAGTGC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
```

RESULT 8

US-10-097-340-284

; Sequence 284, Application US/10097340

; GENERAL INFORMATION:

; APPLICANT: John MONAHAN

```
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGGCTGTTAACTAGGATTGTCGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGCAAGAACAAATGTCACAGGCTGAAATTTATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrlleGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrlleGlnMetLeu 40
QY 320 GAATCCAAACAAATGTGATCAGCTTTCAATGGCTTGGCCAGCTCCAGCTCCAGTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrlleHisThrPhe 60
QY 380 CTTTGGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCAGATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrlleValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATCAAGGATTTTCAAAGATTTTCAAAGATTTGTCGGCCAGTATCTTACACCAAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrlleHisThrArg 100
QY 500 GATGAATCAAGTGGGCTGGGAAAGACATCTCTGAAGAAGATGTGCTAATTTTCATCAAGGTA 559
```

Db 101 AspGluCysLeuThrAlaGlyLysAspLeuLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACGGCTGTGGAAACGGGGGCTTTTCATCAAT 619
Db 121 LeuLysAlaTyAsnGlnThrHisLeuTyAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTAATGGACATCATCTGTAGGACATATTTTAAAGCTGAGAACTCA 679
Db 141 CysThrTyIleGluLeuGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTAATCTGTGAAGTGTGAGTGTATTTATGGGGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTySerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACCTTTGGGCACCAACCAATCAGACAGACAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
QY 860 CTCAATGATCCAAAGTTCATTAGTGCACCTCATCTCAGAGAGTGACAAATCCTGAAGT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGTAAGTCAAAATGCAATAGATGGAGACACTCTGGAAGCT 979
Db 241 AspyLysValTyPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATGTCAGATGACTTTGGAGGCACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyHisArgSerLeuVal 280
QY 1040 AATAATGACACATCTCCAAAGCTCGTCTGATTGCTCAGTGCAGGTCCAAATGGC 1099
Db 281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTGATGAAGTGCAGGATGTATTCCTATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGTGTATGAGTCTCAGTCTCAGTAAACATTTCAAGGATCACCGTG 1219
Db 321 AsnProValValTyIleGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTCTTGTGTCATATGCCACACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyThrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTATCCACGGCAGAACT 1339
Db 361 GlyProAsnTyGlnTyrValProTyThrGlnGlyArgValProTyProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAAAGGACCTTCCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAGTCCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAAATCCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGATTAATATCAATTTACAAATTTGCTGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyThrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAGATGGACATGATGATGTTTATCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTTCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639

Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCACCAATGGAGCTTCCACTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGGTTCAACGGCTGGGTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db 501 LeuTyIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAACGGTGTGAGTGTGTGCTCGCCGAGACCTTACTGTGCTTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyThrCysAlaTyrAspGlySer 540
QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGAGCGCAACAGCAGACAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGCTGACTTACACTTACCATGATATACCATGGCCACACCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGTGTAGAGATAGTACACATTTTGGAAATCAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGTCTATTGCAATTCAGAGGCGAAATGAGAGCGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGCAATTTACTCTGCGCATCGGTGGAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTTAACCTGGAGTCTATTGACACAGAGCATTTGGAAGACTTCTTTCATAAAGT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCCTAATGATGATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGATGAGACTTCATGACGTCAATCAACCCCAATCTCAACACCATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGACCAAGTTTGGAAAGGACCGAAACCAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAAATAATGGAACACTTCAAGAAATAAAGAAAGGTAGAAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCGACCCAGGAGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 9

US-10-247-671-164
; Sequence 164, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671

; CURRENT FILING DATE: 2002-09-18
 ; PRIORITY FILING DATE: 60/323,784
 ; PRIORITY FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 164
 ; LENGTH: 771
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 1930967CD1
 US-10-247-671-164

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-247-671-164 (1-771)

QY 200 ATGGCTGTTAACTAGGATGCTGCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaAlaArgAla 20
 QY 260 AACTATCAGATGGGAAGAACAACTGTCACAGGCTGAATATCTCTCAAGAAATGTTG 319
 Db 21 AsnTyrGlnAsnGlyLeuAsnValProArgLeuLeuSerTyrGlyGluMetLeu 40
 QY 320 GAATCCAACTATGATCACTTTCAATGGCTTGGCCACAGCTCCAGTATTATCATCTTC 379
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGGAACGAGTAGGCTGATGTTGGAGCAAGGATCAATATTTTCATC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLeuAspHisPheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGGATTTTCAAGATGTCGCCAGTATCTTACACCAAGAGA 499
 Db 81 AspLeuValAsnIleLeuAspPheGlnIleValTrpProValSerTyrThrArgArg 100
 QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGCTTAATTTTCATCAAGTA 559
 Db 101 AspGluCysLeuTrpAlaGlyAspIleLeuLeuGlyCysAlaAsnPheIleLeuVal 120
 QY 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAAACGGGGCTTTTCATCCAAAT 619
 Db 121 LeuLeuAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTGGAGACTCA 679
 Db 141 CysThrTyrIleGluIleLeuIleHisProGluAspAsnIlePheLeuLeuGluAsnSer 160
 QY 680 CATTTTGAAACGGCGGTGGGAAGAGTCCATATGACCCCTAAAGCTCTCAGACGATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlySerProTyrAspProLeuLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTAATCTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTGCT 799
 Db 181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCCACTCTGGGCACCCACCATCAGGACAGACGATGATTCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCAAAAGTTTCAATGATGCCACCTCATCTCAGAGAGTGACAATCTCGAAGAT 919
 Db 221 LeuAsnAspProLeuPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGGAAAGACT 979

Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGAAATAGGTAGATATGCAAGATAGCTTTGGAGGCGACAGAGTCTGGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGCAACAACATCTCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCAAAATGCG 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTAGAACTGAGATGATATTCCTTAATGAACCTTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGTATATGAGGTGTTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLeuGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGATGATGTGAGAAGGTGTTCTTGTGTCCATATGCCACACAGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGGCGCAGAACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCACAGAAAACATTTGGTGGTTTGCACCTCTACAAAGGACCTTCTGTATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTTGAAGAAGTATCCAGCCATGATGATCAATCCAGTGTCTTCTATGAACAATGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACGGATGTAATTTCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGCACAGTATGATGTTATGTTTACGAACAGATGTTGGACCGTCTCTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 CTAGTTTCAATCTCTAAGAGAGACTTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
 Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACCTTAAGCAGCAACA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGTTTCAACGGCTGGGTGCTCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
 Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTCGCCCGACACCTTACTGTGCTTGGGATGGTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGAGACGCAAGACCAAGATATTAAGAAAT 1879
 Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
 QY 1880 GGACACCCACTGACTCTGTTACAGTTACCATGATATCACTGATGATGCGCCAGCCCT 1939
 Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAACAGAGAAATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATCAGTCCGAG 1999
 Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGAGCGTGTGCTTATTTGGCAATTCAGAGGCGAAATGAAGAGCGGAAAAAGAG 2059

Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCTTCTCTAGGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATACCTCTGCGCATCGGTGGAAACATGGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACCTCGAAGTCAATGACACAGAGCAATTTGGAGAAGACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGlnHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCAATGATGACATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTCACAGAGCTTCATGTCAGCTCATCAACCAACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTyrTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTCTGCAAGATTTGGAAAAGGACCGAAACCAACGTCGCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGGAAGCAGCTTACAGAAATATAGAAAGGTAGAAACAGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCCACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 10

US-10-262-538-10
; Sequence 10, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)

QY 200 ATGGCTGGTTAACTAGGATCTGTCTTTCTGGGAGTATTACTTACAGCAAGAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGCTGAATATCTCAAGAAATGTGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrThrHisThrPhe 60

QY 380 CTTTTCGATGAGGAACGAGTAGGTGTATCTTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGANTGTGTGGCCAGTATCTTACACGAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATCAATGCAAGTGGCTGGNAAGACATCTGGAAGANTGTCTAATTTTCATCAAGTGA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTGTACCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACCGCCCTGGGAAGTCCATATGACCTAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACGTGATTTTATGGGCGGAGACTTTGCT 799
Db 181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTTGGGCAACCAACCAATCAGGACAGAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTCCGTGAAATCAATAGATGAGAGAACACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTCTAGATATGCAAGAAATGACATTTTGGAGGCGACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCCAATGSC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTGAACTGCAGGATGTATTTCTTAATCAACTTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATAGGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATACATGATGATGTGAGAGGGTGTCTCTTGGTCCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTATCAATGGGTCCCTTATCAAGGAAGAGTCCCTATCCACGGCAGGAACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGAAAAATTGGTGGTTTGTACCTCTCAAGAGGACCTTCTCTGATGATGATTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAGCAAGCTCATCCAGCCATGTACAAATCCAGTGTTCCTTATCAAAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGAGTGTAAATTTATCAATTTTACAAATTTGTCGTAGACCGAGTGGAT 1519

Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
Qy 860 CTCATGATCCAAAGTTCATTAGTCCACCTCATCTCAGAGAGTGAACAATCCCTCAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
Qy 920 GACAAAGTATATCTTTCTTCGTCGAAATGCAATAGATGGAGAACACTCTCGGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
Qy 980 ACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGCCACAGAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
Qy 1040 AATAAATGGACACATCTCTCAAGCTCGTCTGATTTCCTAGTCCAGGTCCCAATGCC 1099
Db 281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
Qy 1100 ATTGACACTCATTTTGATGAACCTGCGAGATGTATTCTTAATGAATTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy 1160 AATCAGTTGTATATGAGTGTATGACTTCCAGTAACTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
Qy 1220 TGTATGTATAGATGATGTGAGAGGTGTTCTTGTGTCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
Qy 1280 GGACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAC 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
Qy 1340 TGTCCAGCAAAACATTTGGTGGTTTGCATCTACAAAGGACCTTCTCGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
Qy 1400 ACCTTTCGAGAGTCCAGCCATGTACATCCAGTGTTCCTATGAACATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTAGACCGAGTGTAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy 1520 CGAAGATGACAGATATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTCTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
Qy 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCTCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
Qy 1640 ACAGTTTTCGGGAACCGACTCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
Qy 1700 CTATATATGTTTCAACCGCTGGGTGTCGCCAGTCCCTTTACACCGGTGTATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
Qy 1760 GGAAGAGGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAGTGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
Qy 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACCCACAAAGACCAAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
Qy 1880 GGAGACCCACTGACTGTTTTCAGACTTACACCATGATATATCAATGCGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580

Qy 1940 GAAGAGAGATCATCTATGTTGTAGAGATAGTACCAATTTTGGATGCAGTCCGAAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
Qy 2000 TCGCAGAGAGCGTGGTCTATTGGCAATTCAGAGGCGAAATGAAGACGAAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAsnGluGluArgLysGluGlu 620
Qy 2060 ATCAGAGTGTATGATCATATCATCAGACAGATCAAGGCCTTCTGCTAGCTAGTCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy 2120 CAGAAGGATTCAGCAATTTACCTCTGCGATGCGGTGCAACATGGGTTCATACAACTCT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
Qy 2180 CTTAAGGTAAACCTGGAGTCAATTCACAGAGCATTTTGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
Qy 2240 GATCATGAGATGCTCTTAAGACCAAGAAATGTCCATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
Qy 2300 GTCTGGTACAGAGACTTTCATGCAGCTCATCAACCCACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProLeuLeuAsnThrMetAspGlu 720
Qy 2360 TTCTGTCAACAAGTTTCGAAAAAGGACCGAAACAAACAGCTCGGCAAGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
Qy 2420 CCAGGGAACAGTAAACATCGAAGCACTTACAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
Qy 2480 ACCCAACCAATTTAGAGGCGACCCAGGAGTGTCT 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 12
US-09-791-537-101922 ; Sequence 101922. Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMEB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 101922
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-101922
Alignment Scores:
Pred. No.: 0 Length: 772
Score: 4057.50 Matches: 740
Percent Similarity: 98.58% Conservativeness: 21
Best Local Similarity: 95.85% Mismatches: 10
Query Match: 82.87% Indels: 1
DB: Gaps: 1
US-09-774-490-1 (1-2709) x US-09-791-537-101922 (1-772)
Qy 200 ATGGCGTGGTTAACTAGAGATTGTCTGCTTTTCTGGGAGTATTACTTACAGCAGAGCA 259
|||||

1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyIleLeuLeuThrAlaAla 20
260 AACTATCAGAAATGGAAGAACAAATGTGCCAAGCGTGAATATATCTCAAGAAATGTTG 319
21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
320 GAATCCAACATGTGATCACATTTCAATGGCTGGCCAAAGCTCCAGTTCATCATCCTTC 379
41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTGGATGAGGAACGAGTAGGTGTATGTGGAGCAAGAGGATCACATATTTTCATTC 439
61 LeuLeuAspGluLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGA 499
81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
500 GATGAATCAAGTGGCTGGAAAGACACATCTGGAAGAAAGATGTCTAAATTCATCAGGTA 559
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
560 CTTAAGGCATATAATCAGACTCACCTGTGACGCTGTGGAAACGGGGCTTTTCATCCAA 619
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
620 TGCACCTACATGAAATTTGGACATCATCTCGAGGCAATATTTTAACTGGAGAACTCA 679
141 CysThrTyrIleGluValGlyHisIleProGluAspAsnIlePheLysLeuGlnAspSer 160
680 CATTTGAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTCTGACAGCATCCCTT 739
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTCTGAACTGACGTGATTTATGGGCGAGACTTTGCT 799
181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAACCTTTGGCCACCACCCCAATCAGGACAGACAGCATGATTCAGAGTGG 859
201 IlePheArgThrLeuGlyHisIleHisProIleArgThrGluGlnHisAspSerArgTrp 220
860 CTCATGATCAAAGTTTCAATGATGCCACCTCATCTCAGAGAGTGACAAATCTCGAAGAT 919
221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGACACTCTGGGAAAGCT 979
241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGATAGTTCAGATATGCAAGAAATGACTTTGGAGGCGACAGAGTCTGTGT 1039
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAAATGGAACAATCTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATGGC 1099
281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTATGAACTCGAGGATCTATTTCTCAATGAACCTTTAAAGATCTTAA 1159
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
1160 AATCAGTGTGATGAGTGTTCACGACTTCCAGTAACTTTCAAGGGATCAGCGGTG 1219
321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGATACATGATGTGAGAAGGTGTTCTTGTGTCATATGCCCCACAGGAT 1279
341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCAACTATCAATGGTCCCTTATCAAGGAGAGTCCCTTATCCACGCCCAGGAACT 1339
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

1340 TGTCCCGACAAAACATTTGGTGGTTTGACTCTCAAGAGCACTTCTCTGATGATGTTATA 1399
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAGAAGTATCCAGCCATGACCAATCCAGTGTTCCTATGAAACAATGCCCA 1459
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAAATTTGCTGAGACCGAGTGAT 1519
421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp 440
1520 GCAGAAGATGCAGATGATGATTTATGTTATTCGGACACAGATGTTGGACCGTCTTTAAA 1579
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480
1640 ACAGTTTTCGGGAACCGACTGCTATTTTACAGCAATGGAGCTTTCACCTAAGCAGCAACA 1699
481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
1700 CTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAACGGTGTGATGATGTTGCTCGCCCGAGACCCCTTACTGTGCTGGATGGTCT 1819
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGTATTTTCCACCTGCAAGAGAGCGCAAGACGACAGATATAGAAT 1879
541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
1880 GGAGACCACTGACTCTACTGTTTCAGACTTA---CACCATGATAATCACCTAGGCCACAGC 1936
561 GlyAspProLeuThrHisCysSerAspLeuGlnHisIleAspAsnHisIleGlyHisSer 580
1937 CCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATGAGTCCG 1996
581 LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
1997 AAGTCGACAGAGCGCTGTTTATTTGGCAATTCAGAGGCGAATGAGACGCAAGAAAGAA 2056
601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluAspArgLysGlu 620
2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAGTCTA 2116
621 GluIleArgValGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
2117 CAACAGAAGATTCAGGCAATTTACTCTGCGATCGGTGGAAACATGGTGTTCATCAAACT 2176
641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
2177 CTTCTTAAGGTAAACCTGGGAAGTCATTGACACAGAGCATTTGGGAAGACTTCTTCATAAA 2236
661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLys 680
2237 GATCATCATGAGATGGCTCTTAAGACCAAGAAATGTCCAAATAGCATGACACCTAGCCAG 2296
681 AspAspAspGlyAspGlySerLysThrLysGluMetSerSerSerMetThrProSerGln 700
2297 AAGTCTGGTACAGAGACTTTCAGCTCATCAACACCCCCCAATCTCAACACAGATGAT 2356
701 LysValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
2357 GAGTCTTGTCACAAAGTTTCGAAAAGGACCGGAAACAAACGTCGGCAAGCCAGGACAT 2416
721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHis 740

```
QY 2417 ACCCCAGGAAACAGTAACAAATGGAGCACTTACAGAAATAGAAAGGTAGAAACAGG 2476
   ::: |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760

QY 2477 AGGACCCACGAATTTGAGAGGCCACCCAGGAGTGTC 2512
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 13
US-09-791-72548
; Sequence 72548, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomic, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 72548
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-72548

Alignment Scores:
Pred. No.: 0 Length: 772
Score: 4035.50 Matches: 737
Percent Similarity: 98.32% Conservative: 22
Best Local Similarity: 95.47% Mismatches: 12
Query Match: 82.42% Indels: 1
DB: 22 Gaps: 1

US-09-774-490-1 (1-2709) x US-09-791-537-72548 (1-772)

QY 200 ATGGCTGTTAACTAGAGATTCTCTTTCTGGGAGTATTACTTACAGCAAGACCA 259
Db 1 MetGlyTrpPheThrGlyLeuAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGAAATGGGAAGCAACATGTGCCAGGCTGAAATATCTCAAGAAATGTTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

QY 320 GAATCCAAATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTGTGGCCAGTAGTATCTTACACCAAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATCAAGTGGGCTCGAAAGACATCTCGAAGAAATGTGCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCCAATT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCTACATTGAATTTGGACATCATCTCAGAGCAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluValGlyHisHisProGluAspAsnIlePheLysLeuGlnAspSer 160

QY 680 CATTTTGAAACGGCGCTGGGAAGAGTCCATATGACCTTAAGCTCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
```

```
QY 740 TTAATAGATGAGAAATTATCTCTGGAACATGACGCTGATTTTATGGGCGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200

QY 800 ATCTTCCGAACCTCTTGGGCGACACACCCCAATCAGGACAGACAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 860 CTCAATGATCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240

QY 920 GACAAGTATATCTTTCTTCCGTAATCAATAGATGAGAGACACTCTCGGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260

QY 980 ACTCACGCTAGATAGTGCAGATATGCAAGATGACTTTTGGAGGCGACAGAAAGTCTGTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAAATGGACAACATTTCTCAAAGCTCGTCTGATTTGCTCAGTGCAGGTCGCAATGGC 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCATTTTGTGATGAACATGACAGATGATTTCTTAATGAACCTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320

QY 1160 AATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACACTTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGATGATAGCATGAGTGTGAGAAAGGTGTTCTTGTCTCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTTATCCAGCGCCAGGAAC 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCCAAGCAAAACATTTGGTGGTGTGACTCTCAAGAGCACTTCTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400

QY 1400 ACCTTTCAGAGAAGTATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420

QY 1460 ATAGTGATCAAAACCGGATGTAATTTATCAATTTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

QY 1520 GCAGAAGATGACAGTATGATGTTATGTTTATCGGAACAGAGTGTGGACCGTCTTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GTAGTTTCAATTCCTAAGAGAGACTTGGTATGATTTAGAAGGTTCTGCTGGAGAAGATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480

QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACCTTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520

QY 1760 GGGAAACGGTGTGCTGAGTGTGCTGCCCGAGACCTTACTGCTGCTGGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
```


1220 TGTATGATATAGCATGATGTGAGAGGGTTCCTTGGTCCATATGCCCCACAGGAT 1279
Db
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGCGCCAGAACT 1339
Db
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCCCGCAAAACACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1399
Db
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
1400 ACCTTTGCAAGAGTCCATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAAATCCGCCA 1459
Db
401 ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
1460 ATATGATCAAAACGGATTAATATCAATTTACAAATTTGTCGTAGACCGAGTGGAT 1519
Db
421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp 440
1520 CGAGAGATGGACATGATCATGTTATGTTTATCGGACAGAGTGTGGGACCGTTCCTAAA 1579
Db
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValIleGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db
461 ValValSerValProLysGluThrTrpHisAspLeuGluGluValLeuLeuGluMet 480
1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCCAATGAAGCAACAA 1699
Db
481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATTTGGTTCACGGCTGGGTGCGCGCTCCCTTTACAGCGGTGATATTTAC 1759
Db
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 CGGAAGCGGTGCTGAGTGTTCCTCGCCGAGACCTTACTGCTGGTGGATGGTCT 1819
Db
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATAAGAAAT 1879
Db
541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
1880 GGAGACCACTGACACTGCTCAGACTTACAC---CATGATAATCAACATGCCACAGC 1936
Db
561 GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisGlyProSer 580
1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAGACATTTTGGAAATGCAAGTCCG 1996
Db
581 LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
1997 AAGTCGACAGAGCGCTGCTATTGGCAATTCAGAGCGCAATTCAGAGCGCAAGAA 2056
Db
601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnArgSerLysArg 620
2057 GAGATCAGATGGATCATCATCATCAGACAGATCAGGCGCTTCTGCTAGCTAGTCTA 2116
Db
621 GluIleArgMetGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
2117 CAACAGAGGATTCAGGCAATTAACCTCTGCCATGGGTGGAAATGCGGTTTCATACAACT 2176
Db
641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
2177 TTCTTTAAGGTAACTCGAAGTCAATGTACACAGAGCAATTCGGAAGAACTTCTTCATAAA 2236
Db
661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLys 680
2237 GATGATGATGGATGGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACTAGCCAG 2296
QY
681 AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln 700
2297 AAGGTCTGTACAGACTTCATGAGCTTCATCAACCAACCCCACTCATCAACAGTGGAT 2356

701 LysValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY
2357 GAGTCTCTGTGAACAAGTTTGAAGAAGGACCGAAGAACCAACCTCGGCAAGGCCAGCAT 2416
Db
721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHis 740
QY
2417 ACCCAGGGAACACTTAACAATGAAGCACCTTACAGAAAAATAAGAAAGGTAGAAACAGG 2476
Db
741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY
2477 AGGACCCACGAATTTGAGGGCCACCCAGGAGTGTCT 2512
Db
761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 15
US-09-791-537-128280
; Sequence 128280, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 128280
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-128280

Alignment Scores:
Pred. No.: 0 Length: 772
Score: 3989.50 Matches: 730
Percent Similarity: 97.28% Conservative: 21
Best Local Similarity: 94.56% Mismatches: 20
Query Match: 81.48% Indels: 1
DB: 22 Gaps: 1

US-09-774-490-1 (1-2709) x US-09-791-537-128280 (1-772)

QY 200 ATGGCTGGTTAACTAGGATTTCTGCTTTTCTGGGAGTATTACTTACAGCAGACGA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGGAAGAACAAATGTGCAAGGCTGAAATTCCTACAAAGAAATGTTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCACACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACTGGTTAATATACAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArg 100
QY 500 GATGAATCAAGTGGCTGGAAAAACATCTCTGAAGAAATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGCATATATACAGACTCCTTGTACCGCTGTGGAAAGGGGCTTTTCATCAATT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

Qy ⁴	620	TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCA	679
Db	141	CysThrTyrlleGluValGlyHisHisProGluAspAsnIlePheLysLeuGlnAspSer	160
Qy	680	CATTTTGAAGACGCGTGGGAAGAGTCCATATCACCTAAGCTGCTCAGACGATCCCTT	739
Db	161	HisPheGluasnGlyArgGlySerProTyrAspProLysLeuLeuThrAlaSerLeu	180
Qy	740	TTAATAGATGGAGAATTATACTCTCGAACTGTCAGCTGATTTTATGGGGCGAGACTTTGCT	799
Db	181	LeuileAspGlyGluLeuTyrSerGlyThrAlaalaAspPheMetGlyArgAspPheAla	200
Qy	800	ATCTTCGCAACTCTTGGCCACCACCCCAATCAGGACAGAGCAGCATGATTCAGGTGG	859
Db	201	IlePheArgThrLeuGlyAspHisHisProIleArgThrGluGlnHisAspSerArgTyr	220
Qy	860	CTCAATGATCCAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGACAATCCTGAAGAT	919
Db	221	LeuasnAspProArgPheIleSerAlaHisLeuIleProGluSerAspMetProGluAsp	240
Qy	920	GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGGAGAACTCTCTGGAAAAAGCT	979
Db	241	AspLysValTyrPhePheArgGluasnAlaileGlyGlyGluHisSerGlyLysAla	260
Qy	980	ACTCACCGTAGAATAGTGCAGATATGCAAGATGACTTTGGAGGGCACAAGAGTCTGGTG	1039
Db	261	ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal	280
Qy	1040	AATAAATGGACAACATCTCCAAAGCTCGTCTGATTTGCTCAGTCCAGGCTCCAAATGGC	1099
Db	281	AsnLysIrrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly	300
Qy	1100	ATTGACACTCATTTTGATGAAGTCCAGGATGTATTCTCTAATGAAGTTTAAAGATCCTAAA	1159
Db	301	IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys	320
Qy	1160	AATCCAGTTGTATGAGAGTGTATTACACTTCAGTAACATTTTCAAGGATCAGCCGTG	1219
Db	321	AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal	340
Qy	1220	TGATGTATAGCATGAGTGTGTCAGAAAGGGTGTCTTGTCATATGCCACAGGAT	1279
Db	341	CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp	360
Qy	1280	GGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAGCT	1339
Db	361	GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr	380
Qy	1340	TGTCCTCAGCAAAACATTTGGTGGTTTGACTCTACAAAGGACCTTCCTGATGATGTTATA	1399
Db	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
Qy	1400	ACCTTTCAGAGAAGTCAATCCAGCCATCACAATCCAGTGTTCCTATGAACAATCCGCCA	1459
Db	401	ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro	420
Qy	1460	ATAGTGATCAAAACCGAGTGAATATCAATTTACACAAATGTCGTAGACCCAGTGCAT	1519
Db	421	IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp	440
Qy	1520	GCAGAAGATGGACAGTATGATGTATTGTTATTCGGAACAGATGTGGGACCGTTCCTAAA	1579
Db	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
Qy	1580	GTAGTTTCAATCTCTAAGGAGACTTGTGTATGATTTAGAAGGTTCTGCTGGAGAAATG	1639
Db	461	ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet	480
Qy	1640	ACAGTTTTTCGGGAACCGACTGCTATTTTTCAGCAATGAGCTTTCCTCACTAAGCAGCAAA	1699
Db	481	ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
Qy	1700	CTATATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGTGATATTAC	1759

Db	501	LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr	520
QY	1760	GGGAACCGTGTGCTAGTGTCTCCCTCGCCGAGACCCCTTACTGTGCTTGGGATGTTCT	1819
Db	521	GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer	540
QY	1820	GCATGTTCTCGCTATTTTCCCACTGCAAGAGAGACGCAAGACGACAAGATATAAGAAAT	1879
Db	541	SerCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn	560
QY	1880	GGAGACCCACTGACTCACTGTTCCAGACTTACAC---CATGATTAATCACATGGCCACAGC	1936
Db	561	GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisHisGlyProSer	580
QY	1937	CCTGAACAGAGAAATCATCTATGTGTAGAGAAATAGTAGCACATTTTGGAAATGCATCTCG	1996
Db	581	LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro	600
QY	1997	AAATCGCAGAGAGCGTGTCTATTGCGCAATTCGACAGCGCAAAATCAAGAGCGAAAGAA	2056
Db	601	LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnArgSerLysArg	620
QY	2057	GAGATCAGATGGATGATCATATCATCAGACAGATCAAGCGCTTCTGTACGTAGTCTA	2116
Db	621	GluIleArgMetGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu	640
QY	2117	CAACAGAAAGTTCAGGCANATTACCTCTGCCATCGGTGGGAACATGGTTCATACAACT	2176
Db	641	GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr	660
QY	2177	CTTCTTAAGGTAAACCTTGAAGTCATTGACACAGACAGCATTTGGAAGAACTTCTTCATAAA	2236
Db	661	LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLys	680
QY	2237	GATGATGATGGAGATGGCTCTAAGACCAAGAATAATGCCAATAGCATGACACCTAGCCAG	2296
Db	681	AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln	700
QY	2297	AAGTCTGTGTACAGACATTCATGAGCTCATCAACCCCATCTCAACAGATGGAT	2356
Db	701	LysValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp	720
QY	2357	GAGTTCGTCAACAAGTTTGGAAAAGGACCGGAACAACCTCGCAAGGCCAGACAT	2416
Db	721	GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHis	740
QY	2417	ACCCGAGGGAACAGTAACAAATGGAAGCACTTCAAGAAAAATGAAGAGGTAGAAACAGG	2476
Db	741	SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg	760
QY	2477	AGGACCCACGAATTTGAGAGGCCACCCAGAGTGTC	2512
Db	761	ArgThrHisGluPheGluArgAlaProArgSerVal	772

Search completed: August 3, 2003, 10:19:46
Job time : 462 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:01:29 ; Search time 12 Seconds
(without alignments)
5659.545 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttattttatcgatg.....agggttttttctcctaacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 67526 seqs, 12534983 residues

Total number of hits satisfying chosen parameters: 135052

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09774490/runat_03082003_095417_25391/app_query.fasta_1.2887
-DB=Pending Patents AA New -OFMT=fastan -SUFFIX=rapn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1.1.9 @runat_03082003_095417_25391 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents AA New:*
1: /cgn2_6/ptodata/2/paa/pct NEW COMB.pcp:*
2: /cgn2_6/ptodata/2/paa/US06 NEW COMB.pcp:*
3: /cgn2_6/ptodata/2/paa/US07 NEW COMB.pcp:*
4: /cgn2_6/ptodata/2/paa/US08 NEW COMB.pcp:*
5: /cgn2_6/ptodata/2/paa/US09 NEW COMB.pcp:*
6: /cgn2_6/ptodata/2/paa/US10 NEW COMB.pcp:*
7: /cgn2_6/ptodata/2/paa/US60 NEW COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #	Score	Match	Length	ID	Description
1	2044	41.7	775	1	PCT-US03-09929-92	Sequence 92, Appl
2	2044	41.7	775	1	PCT-US03-09929-94	Sequence 94, Appl
3	2044	41.7	775	1	PCT-US03-09929-96	Sequence 96, Appl
4	2044	41.7	775	1	PCT-US03-09929-98	Sequence 98, Appl
5	1917	39.2	636	6	US-10-408-765A-2003	Sequence 2003, Ap
6	1302.5	26.6	514	6	US-10-357-820-28	Sequence 28, Appl
7	1136.5	23.2	431	6	US-10-369-072-97	Sequence 97, Appl
8	966	19.7	834	6	US-10-369-072-47	Sequence 47, Appl
9	964	19.7	430	6	US-10-369-072-91	Sequence 91, Appl
10	952.5	19.5	638	1	PCT-US03-09929-10	Sequence 10, Appl
11	952.5	19.5	833	6	US-10-369-072-14	Sequence 14, Appl

12	950.503	19.4	833	1	PCT-US03-09929-6	Sequence 6, Appl
13	950.5	19.4	833	6	US-10-369-072-16	Sequence 16, Appl
14	950.5	19.4	833	6	US-10-369-072-18	Sequence 18, Appl
15	949.503	19.4	634	1	PCT-US03-09929-2	Sequence 2, Appl
16	949.5	19.4	638	1	PCT-US03-09929-8	Sequence 8, Appl
17	949.503	19.4	833	1	PCT-US03-09929-12	Sequence 12, Appl
18	943	19.3	963	6	US-10-369-072-46	Sequence 46, Appl
19	937	19.1	838	6	US-10-369-072-49	Sequence 49, Appl
20	889	18.2	893	6	US-10-369-072-38	Sequence 38, Appl
21	857.5	17.5	782	6	US-10-369-072-39	Sequence 39, Appl
22	856.503	17.5	666	1	PCT-US03-09929-56	Sequence 56, Appl
23	855.5	17.5	878	1	PCT-US03-09929-54	Sequence 54, Appl
24	854.503	17.5	939	1	PCT-US03-09929-32	Sequence 32, Appl
25	854	17.4	640	1	PCT-US03-09929-16	Sequence 16, Appl
26	854	17.4	649	1	PCT-US03-09929-50	Sequence 50, Appl
27	854	17.4	861	1	PCT-US03-09929-52	Sequence 52, Appl
28	854	17.4	1035	1	PCT-US03-09929-18	Sequence 18, Appl
29	854.003	17.4	1047	1	PCT-US03-09929-14	Sequence 14, Appl
30	854.003	17.4	1047	1	PCT-US03-09929-48	Sequence 48, Appl
31	852.503	17.4	626	1	PCT-US03-09929-34	Sequence 34, Appl
32	852.5	17.4	1018	1	PCT-US03-09929-28	Sequence 28, Appl
33	850.503	17.4	998	1	PCT-US03-09929-20	Sequence 20, Appl
34	839.5	17.1	981	1	PCT-US03-09929-30	Sequence 30, Appl
35	834	17.0	756	6	US-10-369-072-8	Sequence 8, Appl
36	827.5	16.9	1017	6	US-10-451-010-2	Sequence 2, Appl
37	825.5	16.9	754	6	US-10-369-072-10	Sequence 10, Appl
38	825.5	16.9	761	6	US-10-293-244-3947	Sequence 3947, Ap
39	825.003	16.9	495	1	PCT-US03-09929-4	Sequence 4, Appl
40	817	16.7	1070	6	US-10-461-791-8	Sequence 8, Appl
41	817	16.7	1070	6	US-10-461-792-8	Sequence 8, Appl
42	817	16.7	1086	6	US-10-461-791-4	Sequence 4, Appl
43	817	16.7	1086	6	US-10-461-792-4	Sequence 4, Appl
44	816.5	16.7	641	6	US-10-461-791-13	Sequence 13, Appl
45	816.5	16.7	641	6	US-10-461-792-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

PCT-US03-09929-92
; Sequence 92, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Curaseq version 0.1
; SEQ ID NO 92
; LENGTH: 775
; TYPE: PRT

ORGANISM: Homo sapiens
PCT-US03-09929-92

Alignment Scores:

Pred. No.: 7,64e-165 Length: 775
Score: 2044.00 Matches: 374
Percent Similarity: 67.57% Conservative: 147
Best Local Similarity: 48.51% Mismatches: 234
Query Match: 41.75% Indels: 16
DB: 1 Gaps: 8

US-09-774-490-1 (1-2709) x PCT-US03-09929-92 (1-775)

QY 218 ATTGCTGCTCTTTCTGGGAGTATCTTACAGCAGAGCAAACTATCAGATCGGAG 277
DB 8 IIEThrLeuLeuLeuTrpGlyTyrLeuLeuGluLeuTrpThrGlyGlyHisThrAlaAsp 27
QY 278 AACAAATGTGCCAGGCTGAATATCTTACAAAGAAATGTTGGAATCCAAATGTGATC 337
DB 28 ThrThrHisProArgLeuArgLeuSerHisLysGluLeuLeuAsnLeuAsnArgThrSer 47
QY 338 ACTTTCATGGCTTGGCCAACAGCTCCAGTTATCATCTCTCTTTGGATGAGAACGG 397
DB 48 IIEPheHisSerProPheGlyPheLeuAspLeuHisThrMetLeuLeuAspGluTyrGln 67
QY 398 ACTAGGCTGTATGTTGGACCAAGGATCACATATTTTCATTCGACCTGTTTAATATCAAG 457
DB 68 GluArgLeuPheValGlyGlyArgAspLeuValTyrSerLeuSerLeuGluArgIleSer 87
QY 458 GAT---TTTCAAAGATTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514
DB 88 AspGlyTyrLysGluIleHisTrpProSerThrAlaLeuLysMetGluGluCysIleMet 107
QY 515 GCTGGAAGAAGACATCCTGGAAGATGTGCTAATTTTCATCAAGGTACTTAAAGGCATATAT 574
DB 108 LysGlyLysAsp---AlaGlyGluCysAlaAsnTyrValArgValLeuHisHisTyrAsn 126
QY 575 CAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCATCCAAATTTGCACTTACATTCGA 634
DB 127 ArgThrHisLeuLeuThrCysGlyThrGlyAlaPheAspProValCysAlaPheIleArg 146
QY 635 ATTGACATCATCTGAGACAAATATTTTAACTGGAGAACTCACATTTTGAACAGCGC 694
DB 147 ValGlyTyrHisLeuGluAspProLeuPheHisLeuGluSerProArgSerGluArgGly 166
QY 695 CQTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATCGAGAA 754
DB 167 ArgGlyArgCysProPheAspProSerSerPheIleSerThrLeuIleGlySerGlu 186
QY 755 TTATCTCTGGAACCTGACCTCATTTTATGGGGCAGAGCTTTGCTATCTTCCGAACTCTT 814
DB 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpSerArgAspAlaIlePheArgSerMet 206
QY 815 GGGCACCACCCCAATCAGACAGACAGCATGATTCAGGTGCTCAATGATCCAAAG 874
DB 207 GlyArgLeuAlaHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 226
QY 875 TTCATTAGTGCCACCTCATCTCAGAGTGATCAATCTGGAAGATGACAAAGTATACTTT 934
DB 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysValTyrPhe 246
QY 935 TTCTTCCGTAATGCAATAGATGGAGAACACTCTGGAAAGCTTACTCACGCTAGAATA 994
DB 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAsnAlaHisAlaIleTyrThrArgVal 266
QY 995 GGTCAATATGCAAGAATGACTTTTGGAGGCGCAGAGTCTCGTGAATAAATGGAACA 1054
DB 267 GlyArgLeuCysValAsnAspValGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
QY 1055 TTCTCTAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATGCGATGCACTCATTTT 1114
DB 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306

QY 1115 GATCAACTGCAGGATGTATTCCTAATGAACATTTAAAGATCCTAAATCCAGTTGTATAT 1174
DB 307 AspGluLeuGluAspValPheLeuLeuProThrArgPheHisLysAsnProValIlePhe 326
QY 1175 GGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGATCAGCGGTGTGTATGTATAGCATG 1234
DB 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaIleCysValTyrHisMet 346
QY 1235 AGTCATGTGAGAAGGGTGTCTTGGTCCATATCCACAGGATGGAACCAACTATCAA 1294
DB 347 SerSerIleArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
QY 1295 TGGTGTGCTTATCAAGGAAGAGTCCCTATCCAGCGGCAGCAACTGTGCCAGCAAAACA 1354
DB 367 TrpSerValTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
QY 1355 TTTGGTGGT---TTTGACTCTACAAGACCTTCTCTGATGATGTATTAACCTTTGCAAGA 1411
DB 387 AsnGlyGlyArgTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
QY 1412 AGTCATCCAGCCATGATCAATCCAGTGTTCCTATGACATCGCCCAATAGTGATCAA 1471
DB 407 SerHisProLeuMetTyrGlnAlaIleLysProAlaHisLysLysProIleuValLys 426
QY 1472 ACGATGTAAATATCAATTTTACAAATTTCTGCTAGACCGAGTGGATGCAAGATCGA 1531
DB 427 ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 446
QY 1532 CAGTATGATGTATGTTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCAATT 1591
DB 447 GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 466
QY 1592 CCTAAGGAGACTTGGTATGATTTAGAAGAGCTTCTGCTGGAGAAATACACAGTTTTCGG 1651
DB 467 TyrAsnGlnGluMetGluSerMetGluGluValIleLeuGluGluLeuGluIlePheLys 486
QY 1652 GAACCGACTGTATTTTCCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATGCT 1711
DB 487 AspProValProIleIleSerMetGluLysSerLysArgGlnGlnLeuTyrIleGly 506
QY 1712 TCAACGGCTGGGTGGCCAGCTCCCTTTACCGGTGTGTATTTTACGGGAAGCGTGT 1771
DB 507 SerAlaSerAlaValAlaGlnValArgPheHisCysAspMetTyrGlySerAlaCys 526
QY 1772 GCTCAGTGTTCCTCGCCGAGACCTTACTGTCTTGGAGTGTCTGTCATGTTCTCGC 1831
DB 527 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546
QY 1832 TATTTTCCCACT-----GCAAAGAGACGCAACAGACCAAGATATAGAATCGA 1882
DB 547 TyrTyrProThrGlyThrHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566
QY 1883 GACCCACTGACTCTGTTTCAGACTTACCATGATATACCATGCGCCAGCCCTGAA 1942
DB 567 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspLysThrGlu 586
QY 1943 GAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGGAACTCAGTCCGAGTCCG 2002
DB 587 GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
QY 2003 CAGAGAGCGCTGTCTATTGGCAATTCAGAGGCGCAATTAAGAGCGCAAGAGATC 2062
DB 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal 626
QY 2063 AGATGATGATCATATCATCAGGACAGATCAAGCCTTCTGCTACGTAGTCTACAAAC 2122
DB 627 LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
QY 2123 AAGGATTCAGCAATTAACCTCTCGCATCGGTGGAACATGGGTTCATCAACAACCTCTCT 2182
DB 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg 666
QY 2183 AAGGTAACCTTGAAGTCAATTGACACAGACATTTGGAAGAACTTCTTCAATAAAGATGAT 2242

Db	88	AspGlyTyrIysGluIleHisTrpProSerThrAlaLeuLysMetGluGluCysIleWet	107
Qy	515	GCTGMAAAGACATCCTCAAAAGAAATGTCTAAATTTCAATCAAGGTACTTAAAGGCATATAAT	574
Db	108	LysGlyIysAsp--AaGlyGluCysAlaAsnTyrValArgValLeuHisIleTyrAsn	126
Qy	575	CAGACTCATCTGTAGCGCTGTGGAAACGGGGCTTTTCATCAATTTTCACCTACATTCGAA	634
Db	127	ArgThrHisLeuLeuThrCysGlyThrGlyAlaPheAspProValCysAlaPheIleArg	146
Qy	635	ATTGGACATCATCTCGAGCAAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAACGGC	694
Db	147	ValGlyTyrHisLeuGluAspProLeuPheHisLeuGluSerProArgSerGluArgGly	166
Qy	695	CGTGGGAAGAGTCCATATGACCCCTAAAGCTGTCGACAGCATCCCTTTTAAATAGATGGAGAA	754
Db	167	ArgGlyArgCysProPheAspProSerSerPheIleSerThrLeuIleGlySerGlu	186
Qy	755	TTTACTCTGGAACCTGCAGCTGATTATTTATGGGGGAGACTTTGTCTATCTTCGGAACCTCT	814
Db	187	LeuPheAlaGlyLeuTyrSerAspTyrTrpSerArgAspAlaAlaIlePheArgSerMet	206
Qy	815	GGGCACACACCAATCAGACAGAGCAGCATGATTCAGGTCGGCTCAATGATCCAAAG	874
Db	207	GlyArgLeuAlaHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys	226
Qy	875	TTTATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCTGAAGATGACAAAGATATCTTT	934
Db	227	PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysValTyrPhe	246
Qy	935	TTCTTCGGTGAATAATGCAATAGATGGAGAACACACTCTCGAAAGCTACTCACGCTAGAAATA	994
Db	247	PhePheThrGluLysAlaLeuGluAlaGluAsnAlaHisAlaIleTyrThrArgVal	266
Qy	995	GCTCAGATATGCAAGAATGACTTTGGAGGGCAGAGAAGTCTGGTGAATAATAATGGAACAA	1054
Db	267	GlyArgLeuCysValAsnAspValGlyGlyGlnArgIleLeuValAsnLysTrpSerThr	286
Qy	1055	TTCTCTCAAGCTCGTCTGTATTTGCTCAGTGCAGCTCCCAATGCGATTCACACTCATTTT	1114
Db	287	PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrThrPhe	306
Qy	1115	GATGAACCTGCAGGATGTATTCCTAATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATAT	1174
Db	307	AspGluLeuGluAspValPheLeuLeuProThrArgAspHisLysAsnProValIlePhe	326
Qy	1175	GGAGTGTTTAGCATCTCCAGTAACATTTTCAAGGATCAGCCGTGTGTATGTATPAGCATG	1234
Db	327	GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaIleCysValTyrHisMet	346
Qy	1235	AGTGATGTGAGAAGGGTTCCTTGGTTCATATGCCACAGGGATGACCCCACTATCAA	1294
Db	347	SerSerIleArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis	366
Qy	1295	TGGTGTGCTTATCAAGGAAGAGTCCCTATCCAGGGCAGAACTCTCCAGCAAAACA	1354
Db	367	TrpSerValTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal	386
Qy	1355	TTTGGTGGT---TTTGACTCTACAAGAGACTTCCTCATGATGATTTAACTTTGCAAGA	1411
Db	387	AsnGlyGlyArgTyrGlyThrLysAspTyrProAspAspAlaIleArgPheAlaArg	406
Qy	1412	AGTCATCCAGCCANGTACATCCAGTCTTCCCTATGAACAACTGCCCAATAGTATGATCAA	1471
Db	407	SerHisProLeuMetTyrGlnAlaIleGlyProAlaHisLysLysProIleLeuValLys	426
Qy	1472	ACGGATGTAAATTAATCAATTTACAAATTTGCTCAGACCGAGTGGATGCAGAAAGATGGA	1531
Db	427	ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValaspArgValGluAlaGluaspGly	446
Qy	1532	CAGTATGATGTATGTTTATCGGAACAGATGTTGGGACCGTCTCTTAAAGTAGTTTCAATT	1591

Db	447	GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle	466
Qy	1592	CCTAAGGAGACTGGTATGATATTTAGAGAGGTTCTGCTGGAAGAAATGACAGATTTTTCGG	1651
Db	467	TyrAsnGlnGluMetGluSerMetGluGluValIleuGluGluGluGlnIlePheLys	486
Qy	1652	GAACCGACTCTATTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAATATATATTGGT	1711
Db	487	AspProValProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly	506
Qy	1712	TCAACGGCTGGGTTGGCCAGCTCCCTTTACACGGGTGTGATATATTTACGGGAAGCGTGT	1771
Db	507	SerAlaSerAlaValAlaGlnValArgPheHisCysAspMetTyrGlySerAlaCys	526
Qy	1772	GCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTTGGGATGTTCTGCATGTTCTCGC	1831
Db	527	AlaaspCysLeuAlaArgAspProTyrCysAlaIrrPaspdylIleSerCysSerArg	546
Qy	1832	TATTTTCCCCT-----GCAAGAGAGCGCACAGACGACAAGATATAAGAAATGGA	1882
Db	547	TyrTyrProThrGlyThrHisAlaLysArgPheArgArgGlnAspValArgHisGly	566
Qy	1883	GACCCACTGACTCCTGTTCCAGACTTACACCATGATAATCACCATGCCACGCCCTGAA	1942
Db	567	AsnAlaalaGlnGlnCysPheGlyGlnPheValGlyAspAlaLeuAspLysThrGlu	586
Qy	1943	GAGAGAATCATCTATGGTGTAGAGAAATAGTAGACACATTTTGGAAATGCAGCTCCGAAGTCG	2002
Db	587	GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer	606
Qy	2003	CAGAGAGCGTGTCTATTGGCAATTCGAGAGCGCAAAATGAAGAGCGAAAGAAGAGATC	2062
Db	607	LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal	626
Qy	2063	AGATGTGATCATCATCATCAGACAGATCAAGGCTTCTGCTAGCTAGTCTTACAACAG	2122
Db	627	LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys	646
Qy	2123	AAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAAATCGGGTTCATACAAACTCTTCTT	2182
Db	647	SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg	666
Qy	2183	AAGGTAAACCTGGAAGTCATGACACAGAGCATTTTCGAGAAGCTTCTTCATAAAGATGAT	2242
Db	667	LysIleThrLeuGluValValGluGluGluLysValGluAspMetPheAsnLysAspAsp	686
Qy	2243	GATGGAGATGGCTCTAAGACCAAA-----GAAATGTCCAATAGCATGACACCTAGCCAG	2296
Db	687	GluGluAspArgHisHisArgMetProCysProAlaGlnSerSerIleSerGlnGlyAla	706
Qy	2297	AAGGTCGTGTACAGAGACTTCATGACGCTCATCAACACCCCACTTCACACGATGGAT	2356
Db	707	LysProTrpTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArgValGlu	726
Qy	2357	GAGTTCGTGAACAAAGTTTGGAAAGGACCGGAAACAAGTCGGCAAGGCCAGACAT	2416
Db	727	GluTyrCysGluLysValTrpCysThrAspArgLysArgLysLysLeuLysMet-----	744
Qy	2417	ACCCAGGCAACAGTAAACAAATGGAGCACTTACAA-----GAAATGAAGAAAGTGAAC	2473
Db	745	-----SerProSerLysTrpLysTyrAlaAsnProGlnGluLysLysLeuArgSer	761
Qy	2474	AGGAGGACCCAGAAATTTGAGAGGCCACCCAGG	2506
Db	762	LysProGluHis-----TyrArgLeuProArg	770

RESULT 4
PCT-US03-09329-98
; Sequence 98, Application PC/TUS0309329
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al
; TITLE OF INVENTION: THERAPEUTIC POLYIP
; FILE REFERENCE: 21402-573B-061


```

Db 527 AlaAspCysLeuAlaArgAspProTyrCysAlaTrpAspGlyLeuSerCysSerArg 546
QY 1832 TATTTTCCCACT-----GCAAGAGACGACCAAGACGACCAAGATATAAGAAATGA 1882
Db 547 TyrTyrProThrGlyThrHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566
QY 1883 GACCCACTGACACACGTTCAGACTTACACCATGATTAATCACCATGGCCACACCCCTGAA 1942
Db 567 AsnAlaAlaGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspLysThrGlu 586
QY 1943 GAGAGATCATCTATGCTAGAGATAGTACGACATTTTGGAAATGGAATGCGATCGAAGTCG 2002
Db 587 GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
QY 2003 CAGAGCCCTGCTGCTATTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAAGAGAGATC 2062
Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal 626
QY 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAACAG 2122
Db 627 LysThrAspAspArgValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
QY 2123 AGGATTCAGGCAATTACCTCTGCCATGCGGTGGACATGCGTTCATACAACTCTTCTT 2182
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg 666
QY 2183 AAGTAAACCTCGAAGTCAATCAGACAGAGCAATTTGGAAAGACTTCTTCATAAAGATGAT 2242
Db 667 LysIleThrLeuGluValValGluGluLysValGluAspMetPheAsnLysAspAsp 686
QY 2243 GATGAGATGGCTCTAAGACCAA-----GAAATGTCCAATAGCATGACACCTAGCCAG 2296
Db 687 GluGluAspArgHisArgMetProCysProAlaGlnSerSerIleSerGlnGlyAla 706
QY 2297 AGGTCTGGTACAGAGACTTCATGAGCTCATCAACCAACCCCACTCAACAGATGGAT 2356
Db 707 LysProTrpTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArgValGlu 726
QY 2357 GAGTCTGTGAACAAGTTGGAAAGGACCGCAACCAACGTCGCAAGGCCAGGACAT 2416
Db 727 GluTyrCysGluLysValTrpCysThrAspArgLysArgLysLysLeuLysMet----- 744
QY 2417 ACCCAGGGAACAGTAACAATGGAAGCACTTACAA-----GAAATAGAAAGGTAGAAAC 2473
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLysLysLeuArgSer 761
QY 2474 AGGAGGCCCAAGATTTGAGAGGCCACCCAGG 2506
Db 762 LysProGluHis-----TyrArgLeuProArg 770

```

RESULT 5

```

US-10-408-765A-2003
; Sequence 2003, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Foin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088, 465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2003
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-408-765A-2003

```

Alignment Scores:
Pred. No.: 3,79e-154 Length: 636
Score: 1917.00 Matches: 359
Percent Similarity: 68.44% Conservative: 103
Best Local Similarity: 53.19% Mismatches: 149
Query Match: 39.15% Indels: 64
DB: 6 Gaps: 9

```

US-09-774-490-1 (1-2709) x US-10-408-765A-2003 (1-636)

```

QY 545 AATTTCATCAAGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGG 604
Db 2 AsnPheValLysLeuLeuHisAlaTyrAsnArgThrHisLeuLeuAlaCysGlyThrGly 21
QY 605 GCTTTTCATCAATTTGCACCTACATTCGAAATGGACATCATCTCGAGGACGACATATTTT 664
Db 22 AlaPheHisProThrCysAlaPheValGluValGlyHisArgAlaGluGluProValLeu 41
QY 665 AAGCTGGAGAACTCACATTTTGAACCGCGTGGAAAGAGTCCATATAGCCCTTAAGCTG 724
Db 42 ArgLeuAspProGlyArgIleGluAspGlyLysSerProTyrAspProArgHis 61
QY 725 CTGACAGCATCCCTTTTAATAGATGGAGAATTATATCTCGAACTGCAGCTGATTTATG 784
Db 62 ArgAlaAlaSerValLeuValGlyGluLeuTyrSerGlyValAlaAlaAspLeuMet 81
QY 785 GGGGAGACTTTGCTATCTTCGAACTCTTGGGACACCAACCCCAATCAGGACAGCAG 844
Db 82 GlyArgAspPheThrIlePheArgSerLeuGlyGlnArgProSerLeuArgThrGluPro 101
QY 845 CATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAT 904
Db 102 HisAspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSer 121
QY 905 GACATCTCTGAAGATGACAAAGTATATCTTTCTTCGCGTGAAGTCAATAGATCGAGAA 964
Db 122 GluAsnProAspAspAspLysIleTyrPhePheArgGluThrAlaValGluAlaAla 141
QY 965 CACTCT---GGAAGACTCTCAGCTAGATAGTCAAGATGATGCAAGATGACTTTGGA 1021
Db 142 ProAlaLeuGlyArgLeuSerValSerArgValGlyGlnIleCysArgAsnAspValGly 161
QY 1022 GGGCAGCAAGCTGTGTGAATAAATGGAACACATTTCTCAAGCTCGCTCATGTTGCTCA 1081
Db 162 GlyGlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSer 181
QY 1082 GTGCCAGGTCCAAATGGCAATTCACATTTTGTATGAACTGCAGGATGTATTCCTAATG 1141
Db 182 ValProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeu 200
QY 1142 AACTTTAAGATCCTAAATCCAGTGTATATCGAGTGTGTATGAGTGTTCAGCTTCAGTAACAT 1201
Db 201 SerSerArgAspHisArgThrProLeuLeuTyrAlaValPheSerThrSerSerSerile 220
QY 1202 TTCAGGAGATCAGCCGTGTGTATGTATAGCATGATGATGTGAGAAGGTGTCCTCTCGT 1261
Db 221 PheGlnGlySerAlaValCysValTyrSerMetAsnAspValArgArgAlaPheLeuGly 240
QY 1262 CCATATGCCACAGGATGGACCCCAACTATCAATTCGGTGGCTTATCAAGGAAGAGTCCCC 1321
Db 241 ProPheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 260
QY 1322 TATCCAGGCCAGGAACTTGTCCAGCAAAACATTTGGTGGTGGTGTGTGACTCTCAAGGAC 1381
Db 261 TyrProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAsp 280
QY 1382 CTTCCTCATGATGTTAATCCTTTGCAAGAGTCCATCCAGCCCATGTCATCCAGTGTGT 1441
Db 281 PheProAspAspValIleGlnPheAlaArgAsnHisProLeuMetTyrAsnSerValLeu 300
QY 1442 CCTATGAACAATCGCCCAATAGTATCAAAACGAGTAAATATCAATTAATTAACAAAT 1501

```


Db 43 GluTyrPheSerLeuSerHisHisProLeuAspTyrArgIleLeuLeuMetAspGluAsp 62
QY 395 CGAGTAGCTGTATGTTGGAGCAAGATCATATTTTCATTCGACCTGGTTATATC 454
Db 63 GlnAspArgIleTyrValGlySerLysAspHisIleLeuSerLeuAsnIleAsnIle 82
QY 455 ---AAGGATTTTCAAAAGATTCTGGCCAGTATCTTACACAGAGAGATCAAGTCAAG 511
Db 83 SerGlnGluAlaLeuSerValPheTyrProAlaSerThrIleLysValGluGluCysLys 102
QY 512 TGGGCTGGAAAAAGACATCTCGAAAGATGTCTTAATTTTCATCAAGCTACTTAAGCATAT 571
Db 103 MetAlaGlyLysAspProThrHisIleGlyCysGlyAsnPheValArgValIleGlnThrPhe 122
QY 572 ATACAGACTCATCTGACCCCTGTGGAGACGGGGCTTTTCATCAATTTGCACTACAT 631
Db 123 AsnArgThrHisLeuTyrValCysGlySerGlyAlaPheSerProValCysThrTyrLeu 142
QY 632 GAAATTGGACATCATCTCGAGCACAATATTTTAAGCTGGAGAACTCACATTTTGAAC 691
Db 143 AsnArgGlyArgSerGluAspGlnValPheMetIleAsp---SerLysCysGluSer 161
QY 692 GGCCGTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCTCTTTTAATAGATGA 751
Db 162 GlyLysGlyArgCysSerPheAsnProAsnValAsnThrValSerValMetIleAsnGlu 181
QY 752 GAATTATATCTGGAACCTGACTGATTTTATGGGCGGAGACTTTGCTATCTTCGAACT 811
Db 182 GluLeuPheSerGlyMetTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgSer 201
QY 812 CTGTTGGCCACCCACCCATCAGCAGCAGCAGCATGATTCAGGTGGCTCAATGATCA 871
Db 202 LeuThrLysArgAsnAlaValArgThrAspGlnHisAsnSerLysTyrTrpLeuSerGluPro 221
QY 872 AAGTTCATTAGTCCCACTCATCTCAGAGAGTGAATCTCGAAGATGACAAAGTATAC 931
Db 222 MetPheValAspAlaHisValIleProAspGlyThrAspProAsnAspAlaLysValTyr 241
QY 932 TTTTCTTCGCGAAATGCAATAGATGAGAGAACACTCTCGAAAGACTACTCACCTAGA 991
Db 242 PhePhePheLysGluLysLeuThrAspAsnAsnArgSerThrLysGlnIleHisSerMet 261
QY 992 ATAGCTCAGATATGCAAGAACTCTTTCGAGGCGCAGAGTCTGGTGGTGAATTAATGGACA 1051
Db 262 IleAlaArgIleCysProAsnAspThrGlyGlyLeuArgSerLeuValAsnLysTyrThr 281
QY 1052 ACATTCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGTCCAGTCCAAATGGCATTCACATCAT 1111
Db 282 ThrPheLeuLysAlaArgLeuValCysSerValThrAspGluAspGlyProGluThrHis 301
QY 1112 TTTGATGAAGTGCAGGATCTATTCCTAATGAATCTTAAAGATCTCTAAAGATCCAGTTGTA 1171
Db 302 PheAspGluLeuGluAspValPheLeuLeuGluThrAspAsnProArgThrThrLeuVal 321
QY 1172 TATGAGTGTTCAGCACTCCAGTAAACATTTTCAAGGATCAGCGTGTGTATGATATAGC 1231
Db 322 TyrGlyIlePheThrThrSerSerValPheLysGlySerAlaValCysValTyrHis 341
QY 1232 ATGAGTGTATGTGAGAAGGTGTCTCTGCTCATATGCCACAGGATGGACCACTAT 1291
Db 342 LeuSerAspIleGlnThrValPheAsnGlyProPheAlaHisLysGluGlyProAsnHis 361
QY 1292 CAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAACCTGTGCCAGCAAA 1351
Db 362 GlnLeuIleSerTyrGlnGlyArgIleProTyrProArgProGlyThrCysProGlyGly 381
QY 1352 ACATTT---GGTGGTGTGACTCTCAAGAGACCTCTCTGATGATGATGATTAACCTTTGCA 1408
Db 382 AlaPheThrProAsnMetArgThrThrLysGluPheProAspAspValValThrPheIle 401
QY 1409 AGAAGTCAATCCAGCCATGTACATCCAGTGTCTCTATGCAATCCGACCAATAGTATC 1468
Db 402 ArgAsnHisProLeuMetTyrAsnSerIleTyrProIleHisLysArgProLeuIleVal 421

QY 1469 AAAACGGAGTAAATTTATCAATTTTACACAAATTTGCTAGACCGAGTGGATGCAGAGAT 1528
Db 422 ArgIleGlyThrAspTyrLysTyrThrLysIleAlaValAspArgValAsnAlaAlaAsp 441
QY 1529 GGACAGTATGATGTATGTTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCA 1588
Db 442 GlyArgTyrHisLeuLeuPheLeuGlyThrAspArgGlyThrValGlnLysValValVal 461
QY 1589 ATTCTTAGGAGACTTGGTATGATTAGACAGAGTTCCTGCTGGAAGAAATCACAGTTTT 1648
Db 462 LeuProThrAsnAsnSerValSer---GlyGluLeuIleLeuGluLeuGluValPhe 480
QY 1649 CGGAACCGACTGCTTATTCAGCAATGAGCTTCCATTAAGCAGCAACAATATATAT 1708
Db 481 LysAsnHisAlaProIleThrThrMetLysIleSerSerLysLysGluGlu----- 497
QY 1709 GGTTCACCGCTGGGTGGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db 498 -----ProLysThrArgCysGluThrTrp 505
RESULT 7
US-10-369-072-97
; Sequence 97, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderma, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Shinkets, Catherine B
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Artificial Sequence

FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Domain search
 ; OTHER INFORMATION: result
 US-10-369-072-97

Alignment Scores:

```

Pred. No.: 2,65e-88 Length: 431
Score: 1136.50 Matches: 229
Percent Similarity: 66.52% Conservative: 69
Best Local Similarity: 51.12% Mismatches: 125
Query Match: 23.21% Indels: 25
DB: 6 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-369-072-97 (1-431)
QY 368 TATCATACCTCTCTTTGGATGAGAAACGAGTAGCTGTATGTTGGAGCAAGATCAC 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 PheArgThrLeuLeuAspGluAspArgGlyArgLeuTyrValGlyAlaArgAsnHis 20
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 428 ATATTTCATTCGACCTGGTGAATATATCAAGGATTTTCAA-----AAGATT 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 ValTyr-----ValLeuAsnLeuGluAspLeuSerGluValLeuAsnLeuLysIle 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 GTGTGCGCCAGTATCTTACACCAAGAGATGAATGCAAGTGGGCTCGAAAAGACATCTG 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 GlyTyrProAlaSerCysGluLysCysGluGluCysAsnMetLysGlyLysSerProLeu 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 AAAGATGTGCTAAATTCATCAAGGTACTTAAGGATATATACAGACTCATCTGTACGCC 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 ThrGluCysThrAsnPheIleArgValLeuGlnAlaTyrAsnAspThrHisLeuTyrVal 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 593 TGTGGAACGGGGCTTTTCATCAATTCACCTACATTTGAATTTGGAGATCATCTCGAG 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CysGlyThrAsnAlaPheAsnProLysCysThrLeuIleAsnLeuGly----- 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 653 GACAAATATTTTAAAGTCGAGAACTCACATTTTGAACCGCGTGGGAAGAGTCCATAT 712
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 ---AspLeuPheSerLeuAspAsnAspAsnGluGluSerGlyCysGlyAspCysProTyr 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 713 GACCTTAAGCTGCTGACGACATCCCTTTTAAATAGATGGAGAAATATATCTTGAAGCTGCA 772
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AspProLeuGlyAsnThrThrSerValLeuValGlyGlyGluLeuTyrSerGlyThrAla 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 GCTGATTTTATCGGCGAGACTTCTGCTATCTTCGAACTCTGGACACCCACCAATC 832
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 IleAspPheSerGlyArgAspProSerIleArgGLeuLeuGlySerHisAspGlyLeu 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 AGCAGACGACGATGATTCACAGTGGCTCAATGATCCAAAGTTCATTAGTCCCACTC 892
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 ArgThrGluPheHisAspSerLysTrpLeuAsnLeuProAsnPheValAspSerTyrPro 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 893 ATCTCAGAGAGTACATCTCGAAGTCAAGTCAAAAGTATATCTTTCTCGTGAATAATGCA 952
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 IleHisTyrValHisSerPheSerAspLysValTyrPhePheArgGluThrAla 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 953 ATAGATGGAGAACTCTCGA---AAAGCTACTCAGCTAGATAGGTGAGTGCAGATGCAAG 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 Val-----GluTyrSerAsnCysLysAlaIleHisSerArgValAlaArgValCysLys 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1010 AATGACTTTGGAGGCGACAGAAGTCTGGTGAATAAATGACAAACATTTCTCAAAGCTCGT 1069
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AsnAspProGlyGlyHisSerTyrLeuGluAsnLysTrpThrPheLeuLysAlaArg 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1070 CTGATTTGCTAGTGCAGGTCCAAATGGCATTGACATCATTTTGTGAACTGACGAGAT 1129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 LeuAsnCysSerIleProGly---GluGlyThrProPheTyrPheAsnGluLeuGlnAla 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1130 GTATTCTTAATCACTTTAAAGATCCTAAAAATCCAGTTGTATATGGAGTGTTCAGACT 1189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 AlaPheLeuLeuProThrAspAsnAspThrAspProValLeuTyrGlyValPheThrThr 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1190 TCCAGTAAATTTTCAAGGGATCAGCGGTGTATGTATAGCATGAGTGTGAGAAAGG 1249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 270 SerSerAsnSerSerAlaGlySerAlaValCysAlaPheSerMetLysAspIleAsnGln 289
QY 1250 GTGTTCTCTTGGTCCATATGCCACAGGATGAGCCCACTATCAATGAGTGGTCCCTTATCAA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 ValPheGluGlyProPheLysHisGln---GlyProAsnSerLysTrpLeuProTyrArg 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1310 GGAAGAGTCCCTATCCACGCGCCAGGAACCTTGTCACCAACAAACATTTGGTGGTTTTCAC 1369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 GlyArgValProTyrProArgProGlyGlnCysProAsnSerSerAsnGly----- 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1370 TCTACAAAGGACCTTCTGATGATGATTATACCTTTGCAAGAGTCAATCCAGCCCATGTAC 1429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 -----AspLeuProAspAspThrLeuAsnPheIleArgCysHisProLeuMetAsp 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1430 AATCCAGTGTTCATGAACAATCGCCCAATAGTAGTCAAAACGAGTGAATATATCAA 1489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 AspValValProProLeuHisAsnValProLeuPheValGlyGlnSerGlyAsnTyrArg 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1490 TTTACACAAATTTGCTAGACCGAGTGGATGCAGAAAGATGGACAG---TATGATGTTATG 1546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 LeuThrSerIleAlaValAspArgValGluAlaGlyAspGlyGlnIleTyrThrValLeu 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1547 TTTATCGAACACAGATGTGGGACCGCTTCTTAAAGTAGTTTCAATCTCTAAGAGACTGG 1606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 PheLeuGlyThrAspAspGlyArgValLeuLysGlnValLeuSerArgSerSer 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1607 TATGATTAGAAGAGTCTCTCGAAGAAATGACAGTTTTCGGGACCGACTGCTATT 1666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 AlaSerTyrGluValValValLeuGluGluSerLeuValPheProAspGlyGluProIle 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1667 TCAGCAATGGAGCTTTCACCTAAG 1690
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GlnAsnMetGluLysSerGlnLys 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

```

US-10-369-072-47
; Sequence 47, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerkhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10369,072
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07

```

US-10-369-072-47

Pred. No.:	8.02e-74	Length:	834
Score:	966.00	Matches:	236
Percent Similarity:	51.35%	Conservative:	107
Best Local Similarity:	35.33%	Mismatches:	289
Query Match:	19.73%	Indels:	56
DB:	6	Gaps:	21

US-09-774-490-1 (1-2709) x US-10-369-072-47 (1-834)

206	TGTTAACTAGGATTCCTGCTCTTTCTGGGGAGTATTACTTACAGCAAGACGAAACTAT	265
Qy		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		
253		
254		
255		
256		
257		
258		
259		
260		
261		
262		
263		
264		
265		
266	CAGAATGGGAAGAACAAATGTGCCAAGGCTGAAATTATCCCTACAAAGAAATGTTGGATCC	325
Qy		
267		
268		
269		
270		
271		
272		
273		
274		
275		
276		
277		
278		
279		
280		
281		
282		
283		
284		
285		
286		
287		
288		
289		
290		
291		
292		
293		
294		
295		
296		
297		
298		
299		
300		
301		
302		
303		
304		
305		
306		
307		
308		
309		
310		
311		
312		
313		
314		
315		
316		
317		
318		
319		
320		
321		
322		
323		
324		
325		
326	AACAATGTGATCATCTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATCTCTCTTTTG	385
Qy		
327		
328		
329		
330		
331		
332		
333		
334		
335		
336		
337		
338		
339		
340		
341		
342		
343		
344		
345		
346		
347		
348		
349		
350		
351		
352		
353		
354		
355		
356		
357		
358		
359		
360		
361		
362		

Qy	917	GATGACAAAGTATACATTTTCTTCCGTGAAAAATGCAATAGATGAGAGAACACACTCTGAAAA	976
Db	232	AspAspLysIleTyrPhePheSerGluArgAlaValGluTyrAspCysTyrSerGlu	251
Qy	977	GCTACTCACGCTAGATAGTTCAGATATGCAAGAATGACTTTGGAGGGCACAGAATCTG	1036
Db	252	GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu	271
Qy	1037	GTGAATAAATGGACAACATTCCTCAAAAGTCGTCTGATTTGCTCAGTGCAGCGCCAGTCAAAAT	1096
Db	272	GlnLysLysTyrThrPheLeuLysAlaArgLeuValCysSerAlaPro	288
Qy	1097	GGCATTGACACTCATTTTGATGAACATGCAGAGTGTATTCCTAATGAACCTTTAAAGATCCT	1156
Db	289	AspTrpLysValTyrPheAsnGlnLeuLysAlaValHisThrLeuArgGlyAlaSerTrp	308
Qy	1157	AAAAATCCAGTGTATATGAGTGGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCC	1216
Db	309	HisAsnThrThrPhePheGlyValPheGlnAlaArgTyrGlyAspMetAspLeuSerAla	328
Qy	1217	GTGTGTATGATAGCTAGTGTATGCAGAAAGCGTGTTCCTGGTCCATATGCCACAGG	1276
Db	329	ValCysGluTyrGlnLeuGluGlnIleGlnValPheGluGlyProTyrLysGluTyr	348
Qy	1277	GATGGACCCAACTATCAATGGTGGCTTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGA	1336
Db	349	SerGluGlnAlaGlnLysTyrAlaArgTyrThrAspProValProSerProArgProGly	368
Qy	1337	ACTTGTCCCAGCAAAACATTT-----GGTGGTGTTTGACTCTACAAAGGACCTTCTGAT	1390
Db	369	SerCysIleAsnAsnTrpHisArgAspAsnGlyTyrThrSerSerLeuGluLeuProAsp	388
Qy	1391	GATGTTATAACCTTTGCAAGAGCTCATCCAGCCATGTCACAATCCAGTGTTCCTATGAAC	1450
Db	389	AsnThrLeuAsnPheIleLysLysHisProLeuMetGluAspGlnValLysProArgLeu	408
Qy	1451	AATCGCCCAATAGTATGATCAAAACGGAGTAAATATCAATTTACACAAATGTGCTAGAC	1510
Db	409	GlyArgProLeuLeuValLysLysAsnThrAsn-----PheThrHisValValAlaAsp	426
Qy	1511	CGAGTGGATCCAGAAAGTGA---CAGTATGATGTTATGTTTATCCGACAGAGTGTGGG	1567
Db	427	ArgValProGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly	446
Qy	1568	ACGGTCTTAAACTAGTTTCAATTCCTAAGGACACTTGGTATGATTTAGAAGAGGTCTG	1627
Db	447	TrpLeuLysAlaValSerLeu-----GlyProTrpIleHis-----Met	460
Qy	1628	CTGGAAGAAATGACAGTTTT---CGGGAACCGACTGCTATTATTCAGCAATGAGCTTCC	1684
Db	461	ValGluGluLeuGlnValPheAspGlnGluPro-----ValGluSerLeuValLeuSer	478
Qy	1685	ACTAAGCAGCAACACTATATATTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACAC	1744
Db	479	GlnSerLysLysValLeuPheAlaGlySerArgSerGlnLeuValGlnLeuSerLeuAla	498
Qy	1745	CGGTGTGATATTACGGGAAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTACTGT	1804
Db	499	AspCysThrLysTyr---ArgPheCysValAspCysValLeuAlaArgAspProTyrCys	517
Qy	1805	GCTTGGGATGGTTCTCGCATGTTCTCGCTATTTTCCCACTGCAGAGAGACGCAACAGCA	1864
Db	518	AlaTrpAsnValAsnThr---SerArgCysValAlaThrThrSerGlyArgSerGlySer	536
Qy	1865	-----CAAGATATAGAANAATGGAGACCCACTGACTCATCTGTTTCAGACTTACACCAT	1915
Db	537	PheLeuValGlnHisValAlaAsnLeuAspThrSerLysMetCys-----	551
Qy	1916	GATAAATCACCATGGCCACAGCCCTGAAGAGAAATC-----ATCTATGGTGTAGAG	1966
Db	552	---AsnGlnTyrGlyIleLysLysValArgSerIleProLysAsnIleThrValValSer	570
Qy	1967	AATAGTAGACATATTTTGGAAATGTCAGTCCGAAAGTCGACAGAGCGCTGCTATTGGCAA	2026

571	GlyThrAspLeuValLeuProCysHisLeuSerAsnLeuAlaHisAlaHisTrpThr	590
2027	TTC-----CAGAGCGCAATGAAGCGCGAAACAAGAGATCAGAGTGGATGATCATC	2080
591	PheGlySerGlnAspLeuProAlaGluGlnProGlySerPheLeuTyArgThrGlyLeu	610
2081	ATCAGGACAGATCAAGCGCTTCCTGCTACGTAGTCTACAACAGAGAGATTGAGCAATTAC	2140
611	-----GlnAlaLeuValMetAlaAlaGlnSerArgHisSerGlyProTyr	626
2141	CTCTGCCATGCGGTGGAACATGGG	2164
627	ArgCysTyrSerGluGluGlnGly	634

RESULT 9

```

US-10-369-072-91
; Sequence 91, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkerez, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 91
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain search
; OTHER INFORMATION: result
US-10-369-072-91
Alignment Scores:

```

```
QY 1313 AGAGTCCCTATCCAGCGCCAGAACTTGTCCAGCAAAACATTTGGTGGTTTGACTCT 1372
Db 310 LeuValProPheProArgProGlyThrCysProAsnAsnSerLeu-----Ser 325
QY 1373 ACAAGGACCTTCTGATGATGTATAACTTTTGCAGGAAGTCAATCCAGCCATGTACAT 1432
Db 326 SerLysAspLeuProAspThrLeuAsnPheIleLysThrHisProLeuMetAspGlu 345
QY 1433 CCAGTGTTCCTATGAAACATCGCCCAATAGTATGATCAAAACCGATGATAATTTCAAT 1492
Db 346 AlaValGlnProLeuThrGlyArgProLeuLeuValLysThrAspSerAsnTyrLeuLeu 365
QY 1493 ACACAAATTTGCTAGACCGAGTGGATGACAGAGATGGACAGTATGATTTATC 1552
Db 366 ThrSerIleAlaValAspArgValArgThrAspGlyGlyAsnTyrThrValLeuPheLeu 385
QY 1553 GGAACAGATGTTGGGACCGTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTATGAT 1612
Db 386 GlyThrSerAspGlyArgIleLeuLysValValLeu-----SerGluSerSerSer 403
QY 1613 TTAGAAGAGTTCCTGCTGGAAGAAATGACAGATTTTTCGGGAACCCGACTGCTATTT 1672
Db 404 SerGluSerValValLeuGluGluIleSerValPheProGlySerProIleSerAsp 423
QY 1673 ATGGAGCTTTCCTACTAAGCAG 1693
Db 424 LeuValIleSerProLysLys 430

RESULT 10
PCT-US03-09929-10
; Sequence 10, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 10
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-10

Alignment Scores:
Pred. No.: 1.02e-72 Length: 638
Score: 952.50 Matches: 230
Percent Similarity: 50.92% Conservative: 103
Best Local Similarity: 35.17% Mismatches: 250
Query Match: 19.45% Indels: 71
DB: 1 Gaps: 19
```

```
US-09-774-490-1 (1-2709) x PCT-US03-09929-10 (1-638)
QY 278 AACAAATGTCAGAGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAACATGTGATC 337
Db 8 AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrValValArgPhe 27
QY 338 ACTTTCATGCTTGGCCCAACAGCTCCAGTTATCATATCATCTTCTTTTGGATGAGAACGG 397
Db 28 SerGlnThrGlyIle-----GlnAspPheLeuThrLeuThrGluProThr 44
QY 398 AGTAGGCTGTATGTCGAGCAAGATCACATATTTTCATTCGACCTGTTATATCAAG 457
Db 45 GlyLeuLeuTyrValGlyAlaArgGluAlaLeuPheAlaPheSerMetGluAlaLeuGlu 64
QY 458 GATTTTCAAAAGATGTCGTGCCAGTATCTTACACCAAGAGATGAATGCAAGTGGGCT 517
Db 65 LeuGlnGlyAlaIleSerTyrPgluAlaProValGluLysLysThrGluCysIleGlnLys 84
QY 518 GGAAGAGATCCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTTAAGGCATATATCAG 577
Db 85 GlyLysAsnSerGlnThrGluCysPheAsnPheIleArgPheLeuGlnProTyrAsnAla 104
QY 578 ACTCACTTGTAGCTGTGGAACCGGGGCTTTTCATCCAAATTTGACCTACATTTGAAAT 637
Db 105 SerHisLeuTyrValCysGlyThrTyrAlaPheGlnProLysCysThrTyrValAsnMet 124
QY 638 GGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACCGCGCT 697
Db 125 -----LeuThrPheThrLeuGluHisGlyGluPheGluAspGlyLys 138
QY 698 GGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATTA 757
Db 139 GlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeuLeuValAspGlyGluLeu 158
QY 758 TACTCTGAACTGACGCTGATTTTATGGGCGAGACTTTGCTATCTTCCGAACTCTTGGG 817
Db 159 TyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleIleLeuArgAsnMetGly 178
QY 818 CACCACCAACCAATCAGGACAGCAGCATGATTCAGCTGGCTCAATGATCCAAAGTTC 877
Db 179 ProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrPheLysGluProHisPhe 197
QY 878 ATTAGTGCCACCTCATCTCAGAGAGTGCACAAATCCT-----GAAGATGACAAAGTA 928
Db 198 ValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAspAspLysVal 217
QY 929 TACTTTTCTTCGCGTGAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCCAGCT 988
Db 218 TyrPhePhePheArgGluArgAlaValGluSerAspCysTyrAlaGluGlnValAla 237
QY 989 AGAATAGTCAAGATATCAAGAATGACCTTTGAGGGGACACAGTCTGCTGTAATAATGG 1048
Db 238 ArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeuGluArgLysTrp 257
QY 1049 ACAACATTTCTCAAAAGCTCGTCTGATTTGCTCAGTGCAGGCTCCAAATGCGACACT 1108
Db 258 ThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro-----AsnTyrGlnLeu 274
QY 1109 CATTGTGATGAACCTGAGGATGTATTCCTAATGAACCTTTAAAGATCTTAAATCCAGTT 1168
Db 275 TyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTyrPheAsnThr 294
QY 1169 GTATATGAGTGTTCAGGACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTATGTAT 1228
Db 295 PheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAlaIleCysGluTyr 314
QY 1229 AGCATGAGTGTGAGAAAGGCTGCTCTTGGTCCATATGCCACAGGATGGACCCCAAC 1288
Db 315 GlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyrHisGluGluAla 334
QY 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGCACTTGTCCACG 1348
Db 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGCACTTGTCCACG 1348
```

```
Db 335 GlnLysTrpAspArgTyrThrAspProValProSerProArgProGlySerCysIleAsn 354
Qy 1349 AAAACATTT-----GGTGGTTTTGACTCTACAAGGACCTTCCTGATCATCTTATAACC 1402
Db 355 AsnTrpHisArgHisGlyTyrThrSerSerLeuGluLeuProAsnIleLeuAsn 374
Qy 1403 TTTGCAAGAGTCCATCCAGCCATGTACATCCAGTGTTCCTTATGAACAATCGCCCAATA 1462
Db 375 PheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrpSerArgProLeu 394
Qy 1463 GTGATCAAAACGGATGTAATATTCAATTTACAAATTTGCTAGACCGAGTGGATGCA 1522
Db 395 LeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAspArgValThrGly 412
Qy 1523 GAAGATGGA---CAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCCCTTCTTAA 1579
Db 413 LeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGlyTrpLeuLeuLys 432
Qy 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCCTGCTGGAAGAAATG 1639
Db 433 AlaValSerLeu-----GlyProTrpValHis-----LeuIleGluGluLeu 446
Qy 1640 ACAGTTT---CGGGAACCGACTGCTATTTCAGCAATGAGACTTTCACCTAAGCAGAA 1696
Db 447 GlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSerGlnSerLysLys 464
Qy 1697 CAACATATATTGGTTCACAGCGTGGGTGCTGCCAGCTCCCTTTACACCGGTGTGATTT 1756
Db 465 LeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAlaAlaAspCysMetLys 484
Qy 1757 TACGGAAACCGTGTCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTGGTGGATGT 1816
Db 485 Tyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCysAlaTrpSerVal 503
Qy 1817 TGTGCATGTTCTCGCTATTTCCTCCACTGCAAGAGACGCAAGACGA----- 1864
Db 504 Asn-----ThrSerArgCysValAlaVal 511
Qy 1865 -----CAAGATATAAGAAATGGAGACCCACTGACTCAC 1897
Db 512 GlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAspThrSerGlyIle 531
Qy 1898 TGTTCAGACTTACACCATGATTAATCACCATGCCACAGC-----CCTCAAGAGAGA 1948
Db 532 Cys-----AsnLeuArgGlySerLysLysValArgProThrProLys 545
Qy 1949 ATCATCTATGTTGAGAAATAGTAGCACATTTTGGAAATGAGTCCGAGTCCGAGAGA 2008
Db 546 AsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeu 565
Qy 2009 GCGCTGCTATTGGCAATTCCAGAGCGCAAT-----GAAGAGCGAAAGAGAGATC 2062
Db 566 AlaHisAlaArgTrpThrPheGlyArgAspLeuProAlaGluGlnProGlySerPhe 585
Qy 2063 AGAGTGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGTCTACACAG 2122
Db 586 LeuTyrAspAlaArgLeu-----GlnAlaLeuValValMetAlaAlaGlnPro 601
Qy 2123 AAGATTTCAGCAATTTACTCTGCCATGCGGTGGAAACATGGG 2164
Db 602 ArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 615
RESULT 11
US-10-369-072-14
; Sequence 14, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
```

```
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-14
Alignment Scores:
Pred. No.: 1,11e-72 Length: 833
Score: 952.50 Matches: 236
Percent Similarity: 50.82% Conservative: 106
Best Local Similarity: 35.07% Mismatches: 264
Query Match: 19.45% Indels: 67
DB: 6 Gaps: 22
US-09-774-490-1 (1-2709) x US-10-369-072-14 (1-833)
Qy 206 TGGTTAACTAGATGTCTGCTTTCTGGGGAGTATTACTTACAGCAAGAGCAACTAT 265
Db 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyAlaGluValTrpTrp 25
Qy 266 CAGATCGGAAGACAAATGTCCAGGCTGAAATTCCTACAAAGAAATGTTGGAATCC 325
Db 26 -----AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrVal 41
Qy 326 AACAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTCATCATCTTCTTTG 385
Db 42 ValArgArgPheSerGlnThrGlyIle-----GlnAspPheLeuThrLeuThrLeu 58
Qy 386 GATGAGAACCGAGTAGGCTGTATGTGGACGAAGAGATCATATTTCATTTCGACCTG 445
Db 59 ThrGluProThrGlyLeuLeuTyrValGlyAlaArgGluAlaLeuPheAlaPheSerVal 78
Qy 446 GTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAGAGAGATGAA 505
Db 79 GluAlaLeuGluLeuGlnGlyAlaIleSerTrpGluAlaProValGluLysThrGlu 98
```

```
QY 506 TCGAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAATTTCAATCAAGTACTTAAAG 565
Db 99 CysIleGlnLysGlyLysAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
QY 566 GCATATAATCAGACTCAGCTGTGACGCTGTGGAAGGGGCTTTTCATCAATTTGCACC 625
Db 119 ProTyrAsnAlaSerHisIleuTyrValCysGlyThrTyrAlaPheGlnProLysCysThr 138
QY 626 TCATTTGAAATGGACATCATCTGTAGGACAAATATTTTAAAGCTGGAGAACATTCATTT 685
Db 139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
QY 686 GAAACGGCTGGGAGAGTCCATATGACCCCTAAGCTCTGACAGCATCCCTTTTAATA 745
Db 153 GluAspGlyLeuGlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeuLeuVal 172
QY 746 GATGAGAAATATATCTCTGAACTGCAGCTGATTTTATGGGCGAGACTTCTCATCTTC 805
Db 173 AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeuLeu 192
QY 806 CGAATCTTTGGCACCCACCAATCAGGACAGCAGCATGATTCAGGCTGGCTCAAT 865
Db 193 ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrLeuAsn 211
QY 866 GATCCAAATTCATAGTCCACCATCTCATCTCAGAGAGTGACAACTCT-----GAA 916
Db 212 GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
QY 917 GATGACAAATGATCTTTCTTCGCGTGAATGCAATAGATAGGAGGACACTCTGGAAA 976
Db 232 AspAspLysValTyrPheLeuPheArgGluArgAlaValGluSerAspCysTyrAlaGlu 251
QY 977 GCTACTCAGCTAGATAGTGCAGATATCAAGAATGACTTTGGAGGGCACAGAACTGTG 1036
Db 252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
QY 1037 GTGAATAATGACAACTCTCTCAAGCTGCTGATTTGCTGATTTGCTGATGCTGACGTC 1096
Db 272 GlnArgLysTrpThrPheLeuLysAlaArgLeuAlaCysSerAlaPro-----288
QY 1097 GGCATTGACACTTCTTATGAACTGCAGAGTATCTTCTTAATGAACCTTTAAAGATCCT 1156
Db 289 AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
QY 1157 AAAAAATCCAGTTGTATATGAGTGTATTCAGACTTCCAGTAAACATTTTCAAGGAGTAC 1216
Db 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
QY 1217 GTGTGTATGATAGATGATGATGAGAGGGTGTCTTGTGCTCCATATGCCCCACAGG 1276
Db 329 IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
QY 1277 GATGACCCCAATCATCAATGGTGGCTTATCAAGNAGAGTCCCTATCCAGGCGCAGA 1336
Db 349 HisGluGluAlaGlnLysTrpAspArgfyrThrAspProValProSerProArgProGly 368
QY 1337 ACTGTGCCCAAAAAACATTT-----GGTGTGTTTGACTCTACAAAGACCTTCTCTGAT 1390
Db 369 SerCysIleAsnAsnTrpHisArgArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
QY 1391 GATGTTATACCTTTTGCAGAAGATCATCCAGCCATGTAACATCCAGTGTTCCTATGAAC 1450
Db 389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
QY 1451 AATCCCAATAGTATCAAAACGGATGTAAATTCATTTTACACAAATGTCTGAC 1510
Db 409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
QY 1511 CGAGTGATGACAGAGATGGA---CAGTATGATGTATGTTTATTCGGAACAGATGTGGG 1567
Db 427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyGlnAla 446
QY 1568 ACCGTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAGAGGTTCTG 1627
```

RESULT 12

PCT-US03-09929-6

; Sequence 6, Application PC/TUS0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-573B-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/402,207

; PRIOR FILING DATE: 2002-08-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: CuraSeqList version 0.1
 ; SEQ ID NO 6
 ; LENGTH: 833
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-09929-6

Alignment Scores:
 Pred. No.: 1.63e-72 Length: 833
 Score: 950.50 Matches: 235
 Percent Similarity: 50.00% Conservativity: 104
 Best Local Similarity: 34.66% Mismatches: 262
 Query Match: 19.41% Indels: 77
 DB: 1 Gaps: 21

US-09-774-490-1 (1-2709) x PCT-US03-09929-6 (1-833)

QY 206 TGGTTAACTAGGATTCTGCTCTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTAT 265
 DB 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyLeuGlyAlaGluValTrpTrp 25
 QY 266 CAGAAATGGAAGAACAATGTGCCAGGCTGAATTTATCTCACAAGAATAATGTGGAATCC 325
 DB 26 -----AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrVal 41
 QY 326 AACATGTGATCATCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTCTCTTTG 385
 DB 42 ValArgArgPheSerGlnThrGlyLe-----GlnAspPheLeuThrLeuThrLeu 58
 QY 386 GATGAGGAACGAGTAGGCTGTGTGGAGCAAGAGTACATATTTTTCATTCGACCTG 445
 DB 59 ThrGluProThrGlyLeuLeuLeuValGlyAlaArgAspHisAlaSerAlaLeuGlyVal 78
 QY 446 GTTAATATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGAGATGAA 505
 DB 79 ProValLeuLeuGlnAlaValIleSerTrpGluAlaProValGluLysLysThrGlu 98
 QY 506 TCGAAGTGGCTGGAAAACACATCTGAAAGATGTGCTAATTTTCATCAAGTACTTAAG 565
 DB 99 CysIleGlnLysGlyLysAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
 QY 566 GCATATAATCAGATCAGCTTGTGAGCTGTGGAACGGGGCTTTTTCATCCAAATTTGCACC 625
 DB 119 ProTyrAsnAlaSerHisLeuLeuValCysGlyThrTyrAlaPheGlnProLysCysThr 138
 QY 626 TCATTTGAAATTTGACATCATCTGTAGGACAATATTTTAACTGGAGAACTCACTTTT 685
 DB 139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
 QY 686 GAAACGGCTGGGAGAGTCCATATGACCTTAAGCTGCTCAGACATCCCTTTTAATA 745
 DB 153 GluAspGlyLysGlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeuLeuVal 172
 QY 746 GATGAGAAATTAATCTCTGGAATCTGAGCTGATTTTATGGCGAGACTTCTCTATCTTC 805
 DB 173 AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeuLeu 192
 QY 806 CGAACTCTTGGGCACACCAATCAGGACAGAGCATGATTTCCAGGTGGCTCAAT 865
 DB 193 ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheTrpLeuAsn 211
 QY 866 GATCCAAAGTTTATAGTCCCACTCATCTCAGAGAGTGACAATCCT-----GAA 916
 DB 212 GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
 QY 917 GATGACAAAGTACTTTTCTTCCGTGAAATGCAATAGATGAGAGCACTCTCGAAA 976
 DB 232 AspAspLysValTyrPhePhePheArgGluArgAlaValGluSerAspCysTyrAlaGlu 251
 QY 977 GCTACTCAGCTAGATAGGTGAGATATCAAGATGACTTTGGAGGGCAGCAAGTCTG 1036
 DB 252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271

QY 1037 GTGAATAATGCAACAATCTCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCCAAAT 1096
 DB 272 GlnArgLysTrpThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
 QY 1097 GGCATTGACACTCAATTTTGTGAAGTGTGATTTCTTAATGAACCTTTAAAGATCCT 1156
 DB 289 AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
 QY 1157 AAAAATCCAGTTGATATGAGAGTGTTCAGACTCTCCAGTAACATTTTCAAGGGATCAGCC 1216
 DB 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
 QY 1217 GTGTGTATATACATGATGATGTGAGAGGGTGTTCCTTGGTCCATATGATGCCACAGG 1276
 DB 329 IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
 QY 1277 GATGACCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGA 1336
 DB 349 HisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSerProArgProGly 368
 QY 1337 ACTTGTCACCAAAAACATTT-----GGTGGTTTGTACTCTACAAAGGACCTTCTGAT 1390
 DB 369 SerCysIleAsnAsnTrpHisArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
 QY 1391 GATCTTAACTTTTCAAGAGAGTCCATCCAGCCATGTACATCCAGTGTTCCTATGAAC 1450
 DB 389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
 QY 1451 AATGCCCAATAGTATGATCAAAACCGATGTAATATCAATTTTACACAAATTTGCTAGAC 1510
 DB 409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
 QY 1511 CGATGAGATGAGAGATGGA---CAGTATGATGTTATGTTTATCGGAACAGATGTTGGG 1567
 DB 427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
 QY 1568 ACCGTTCTTAAGTAGTTCATTTTCAATTCCTAAGCAGACTTGGTATGTTTAGAAGAGTCTG 1627
 DB 447 TrpLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
 QY 1628 CTGGAAGAAATGACAGTTTTT---CGGAAACCGACTGCTATTTTACGAAATGAGAGTTC 1684
 DB 461 IleGluGluLeuGlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSer 478
 QY 1685 ACTAAGCAGCAACATATATATTTGTTTCAAGCTGGGTGGGTGGCCACTCCCTTTACAC 1744
 DB 479 GlnSerLysLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
 QY 1745 CGGTGTGATATTTACGGAAAGCGTGTGCTCAGTGTTCCTCGCCGAGACCTTACTGT 1804
 DB 499 AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCys 517
 QY 1805 CTTTGGATGTTCTGCTATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAAGACGA 1864
 DB 518 AlaTrpSerValAsn-----ThrSerArg 525
 QY 1865 -----CAAGTATATAAGAAATGGAGAC 1885
 DB 526 CysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAsp 545
 QY 1886 CCATGACTCACTGTTTCAGACTTACCATCATATCACCATTGCGCCACAGC----- 1936
 DB 546 ThrSerGlyIleCys-----AsnLeuArgGlySerLysLysValArg 559
 QY 1937 CCTGAAGAGAGAAATCATCTATGTTAGAGAAATAGTAGACATTTTGGAAATGCAAGTCCG 1996
 DB 560 ProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeu 579
 QY 1997 AAGTCGCGAGAGCGCTGCTCTATTTGCAATTTCCAGAGGCCAAAT-----GAAGAGCGA 2050
 DB 580 SerSerAsnLeuAlaHisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGluGln 599

QY 2051 AAGAGAGATCAGAGTGGATGATCATATCATCATCAGCAGATCAAGGCTTCTGTACGT 2110
Db 600 ProGlySerPheLeuTyAspAlaArgLeu- - - - -GlnAlaLeuValMet 615
QY 2111 AGCTTACACAGCAAGATTACGGCAATACCTCTGCATGGCGGTGGAACATGGG 2164
Db 616 AlaAlaGlnProArgHisAlaGlyAlaTyHisCysPheSerGluGluGlnGly 633

RESULT 13
US-10-369-072-16
; Sequence 16, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernav, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-16

Alignment Scores:
Pred. No.: 1.63e-72 Length: 833
Score: 950.50 Matches: 235
Percent Similarity: 50.00% Conservative: 104
Best Local Similarity: 34.66% Mismatches: 262
Query Match: 19.41% Indels: 77
DB: 6 Gaps: 21

US-09-774-490-1 (1-2709) x US-10-369-072-16 (1-833)
QY 206 TGGTTAACTAGGATGTCTGCTCTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTAT 265

Db 8 TrpLeu- - - - -LeuAlaAlaArgLeuTrpGlyLeuGlyAlaGluValTrpTrp 25
QY 266 CAGAAATGGGAAGAAACAATGCCAGGCTGAAATTTATCTCTACAAAGAAATGTTGGAATCC 325
Db 26 - - - - -AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrVal 41
QY 326 AACAAATGTGATCACTTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATCATCTTCCTTTG 385
Db 42 ValArgArgPheSerGlnThrGlyLe- - - - -GlnAspPheLeuThrLeuThrLeu 58
QY 386 GATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTTCATTCACGCTG 445
Db 59 ThrGluProThrGlyLeuLeuTyValGlyAlaArgAspHisAlaSerAlaLeuGlyVal 78
QY 446 GTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 79 ProValLeuLeuGlnAlaValIleSerTrpGluAlaProValGluLysThrGlu 98
QY 506 TGCAAGTGGGCTGGAAGACACATCCTGAAAGAATGTGCTAAATTTTCATCAAGGTACTTAAG 565
Db 99 CysIleGlnLysGlyLysAsnAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
QY 566 GCATATAATCAGACTCCTTGTACGCTGTGGAACGGGGCTTTTCATCCTCAATTTGCACC 625
Db 119 ProTyAsnAlaSerHisLeuTyValCysGlyThrTyAlaPheGlnProLysCysThr 138
QY 626 TACATTGAAATGGACATCATCTCGAGGACAATATTTTAAAGCTGGAGCACTCACATTTT 685
Db 139 TyrValAsnMet- - - - -LeuThrPheThrLeuGluHisGlyGluPhe 152
QY 686 GAAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
Db 153 GluAspGlyLysGlyLysCysProTyAspProAlaLysGlyHisAlaGlyLeuLeuVal 172
QY 746 GATGAGAAATATATCTCTGGAACCTGACGCTGATTTTATGGGCGAGACTTGTCTATCTTC 805
Db 173 AspGlyGluLeuTySerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeu 192
QY 806 CGAACTCTTGGGCAACCCACCTCAGGAGAGTGCATGATGATTCAGGCTGGCTCAAT 865
Db 193 ArgAsnMetGlyProHisHisSerMetLysThrGlu- - - - -TyrLeuAlaPheTrpLeuAsn 211
QY 866 GATCAAAAGTTTCATTAGTGCCCACTCATCTCAGAGAGTGACAATCTCT- - - - -GAA 916
Db 212 GluProHisPheValGlySerAlaTyValProGluSerValGlySerPheThrClyAsp 231
QY 917 GATGACAAAGTATATCTTTTCTTCGTAATAATGCAATAGATGGAGAACACTCTGGAATA 976
Db 232 AspAspLysValTyPhePheArgGluAlaValGluSerAspCysTyAlaGlu 251
QY 977 GCTACTCAGCTAGATAGGTACATATGCAAGTATGCAATCTTGGAGGACAGAGTCTG 1036
Db 252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
QY 1037 GTGAATAATGGACAACATTTCTCAAGCTCGTCTGATTTCTCAGTCCAGGTCCTCAAT 1096
Db 272 GlnArgLysTrpThrPheLeuLysAlaArgLeuAlaCysSerAlaPro- - - - - 288
QY 1097 GGCATTGACACTCATTTTGTATGAAGCTGATGATGATTTCTTAATGAACCTTTAAAGATCCT 1156
Db 289 AsnTrpGlnLeuTyPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAsnThrSerTrp 308
QY 1157 AAAAATCCAGTTGATATGAGTGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCC 1216
Db 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyLeuSerAla 328
QY 1217 GTGTGTATGATAGCATGATGATGAGAGGGGTCTTCTGCTGCTCATATGCCACAGG 1276
Db 329 IleCysGluTyGlnLeuGluGluIleGlnArgValPheGluGlyProTyLysGluTy 348
QY 1277 GATGGACCAACTATCATCTGCTGCTTATCAAGGAGAGTCCCTTATCCACGCCACGGA 1336

506 TGCAAGTGGCTGGAAAGACATCTCTGAAGATGTGCTAATTTTCATCAAGTACTTAAG 565
99 CysIleGlnLysGlyLeuAsnGlnThrGluCysPheAsnPhelIleArgPheLeuGln 118
566 GCATATAATCAGACTCAGCTTGTACCGCTGTGAACGGGGCTTTTCATCCAAATTTGCACC 625
119 ProTyrAsnAlaSerHisLeuTyrValCysGlyThrTyrAlaPheGlnProLysCysThr 138
626 TACATTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCAATTT 685
139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
686 GAACACGGCTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGACATCCCTTTTATA 745
153 GluAspGlyLysGlyLeuPheCysProTyrAspProAlaLysGlyHisAlaGlyLeuVal 172
746 GATGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTTGCTATCTTC 805
173 AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeu 192
806 CGAATCTTGGCCACCAACCCATCAGGACAGACGATGATTCAGAGTGGCTCAAT 865
193 ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrLeuAsn 211
866 GATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGAGAGTGACATCT-----GAA 916
212 GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
917 GATGACAAAGTATATCTTTCTTCGTCGAAATGCAATAGATGGAGAACACTCTGGAAGA 976
232 AspLeuValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 251
977 GCTACTACGCTAGATAGTGCAGATGCAAGATGACTTTTGGAGGCGACAGAAAGTCTG 1036
252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
1037 GTGAATAATGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAT 1096
272 GlnArgLysThrThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
1097 GGCATTGACACTCATTTTGTGAATGCTGAGATGATTTCTTAATGAACCTTTAAAGATCCT 1156
289 AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
1157 AAAATCCAGTTGATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 1216
309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
1217 GTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
329 IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
1277 GATGACCAACTATCATATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 1336
349 HisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSerProArgProGly 368
1337 ACTTGTCCTCCCAAAACATTT-----GGTGGTTTGTACTTACAAAGACCTCTCTGAT 1390
369 SerCysIleAsnAsnTrpHisArgArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
1391 GATGTTATACTTTGCAAGAAGTATCCAGCCATGATCAATTCAGTGTGTTTCTTATGAAC 1450
389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
1451 AATGCCCAATAGTATCAAAACGGATGTAATATCAATTTACAAATTTGCTGATAGC 1510
409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
1511 CGAGTGGATGAGAGAGTGA---CAGTATGATGTTATGTTTATCGGAACAGAGTGTGGG 1567
427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
1568 ACCGTTCTTAAAGTAGTTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 1627

447 TrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
1628 CTGGAAGAAATGACAGATTTT---CGGAACCGACTGCTATTTTTCAGCAATGAGCTTTCC 1684
461 IleGluGluLeuGlnLeuPheAspGlnIlePro-----MetArgSerLeuValLeuSer 478
1685 ACTAAGCAGCAACAACTATATATTGTTCAACGGCTGGGGTGGCCAGCTCCCTTTACAC 1744
479 GlnSerLysLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
1745 CGGTGTGATATTTACGGGAAGCGTGTCTCAGTGTGCTCCGCGGAGACCTTACTGT 1804
499 AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCys 517
1805 GCTTGGGATGTTCTGCTGATGTTCTCGCTATTTTCCACTGCAAGAGACGACAGACGA 1864
518 AlaTrpSerValAsn-----ThSerArg 525
1865 -----CAAGATATAGAATGGAGAC 1885
526 CysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAsp 545
1886 CCACTGACTACTGTTTACAGCTTACCATCATATATCACCATGCGCCACAGC----- 1936
546 ThrSerGlyIleCys-----AsnLeuArgGlySerLysLysValArg 559
1937 CCTCAAGAGAGAACTATCTATGTTGTAGAGAATAGTACACATTTTGGAAATGAGTCCG 1996
560 ProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeu 579
1997 AAGTCGACAGAGCGCTGCTTATTTGGCAATTCAGAGCGGCAAT-----GAAGAGCGA 2050
580 SerSerAsnLeuAlaHisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGln 599
2051 AAAGAAGAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTCTAGCT 2110
600 ProGlySerPheLeuTyrAspAlaArgLeu-----GlnAlaLeuValValMet 615
2111 AGTCTACAACAAGAGGATTCAGGCAATTTACCTCTGCCATCGGCGGAACTGGG 2164
616 AlaAlaGlnProArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 633

RESULT 15

PCT-US03-09929-2
; Sequence 2, Application PC/JUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curogen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: CuraseqList version 0.1
 ; SEQ ID NO 2
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-09929-2

Alignment Scores:

Pred. No.: 1.82e-72 Length: 634
 Score: 949.50 Matches: 230
 Percent Similarity: 50.76% Conservative: 102
 Best Local Similarity: 35.17% Mismatches: 251
 Query Match: 19.39% Indels: 71
 DB: 1 Gaps: 19

US-09-774-490-1 (1-2709) x PCT-US03-09929-2 (1-634)

QY 278 AACAAATGCGCAAGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAACAATGTGATC 337
 DB 6 AenLeuValProArgLysThrValSerSerGlyGluLeuAlaThrValValArgPhe 25
 QY 338 ACTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTCTCTTTGGATGAGGAACGG 397
 DB 26 SerGlnThrGlyLeu-----GlnAspPheLeuThrLeuThrGluProThr 42
 QY 398 AGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATCGACTGGTTAATATCAAG 457
 DB 43 GlyLeuLeuTyValGlyAlaArgGluAlaLeuPheAlaPheSerMetGluAlaLeuGlu 62
 QY 458 GATTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAGATGAATGCAAGTGGGCT 517
 DB 63 LeuGlnGlyAlaLeuSerTrpGluAlaProValGluGluLeuThrGluCysIleGlnLys 82
 QY 518 GGAAAGACATCCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTTAAGGCATATAATCAG 577
 DB 83 GlyLysAsnAsnGlnThrGluCysPheAsnPhelIleArgPheLeuGlnProTrpAsnAla 102
 QY 578 ACTGACTGTAGCCCTGGAGCGGGGCTTTTCATCCAAATTTGACCTACATCAATTAAT 637
 DB 103 SerHisLeuTyValCysGlyThrTyAlaPheGlnProLysCysThrTyValAsnMet 122
 QY 638 GGACATCATCCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGT 697
 DB 123 -----LeuThrPheThrLeuGluHisGlyGluPheGluAspGlyLys 136
 QY 698 GGGAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATTA 757
 DB 137 GlyLysCysProTyAspProAlaLysGlyHisAlaGlyLeuLeuValAspGlyGluLeu 156
 QY 758 TACTCTGGAAGTGCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGAACTCTTGGG 817
 DB 157 TyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleIleLeuArgAsnMetGly 176
 QY 818 CACCACCACTCCATCAGGACAGCAGCATGATTCAGGTGCTCAATGATCCAAAGTTC 877
 DB 177 ProHisHisSerMetLysThrGlu---TyrLeuAlaPheTrpLeuAsnGluProHisPhe 195
 QY 878 ATTAGTGCCCACTCATCTCAGAGATGACAACTCT-----GAAGATGACAAAGTA 928
 DB 196 ValGlySerAlaTyValProGluSerValGlySerPheThrGlyAspAspLysVal 215
 QY 929 TACTTTTCTTCGTTGAAATGCATAGATGAGAACACTCTGGAAAGCTACTACCGCT 988
 DB 216 TyrPhePhePheArgGluArgAlaValGluSerAspCysTyAlaGluGlnValValAla 235
 QY 989 AGAATAGGTGCATATGCAAGAAATGACATTTTGAGGCGCACAGAAAGTCTGTGATAAATGG 1048
 DB 236 ArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeuGlnArgLysTrp 255
 QY 1049 ACAACATCTCTCAAGCTGCTGATTTGCTCAGTGCCAGGCTCCAAATGGCATTCACACT 1108
 DB 255 ThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro-----AsnTrpGlnLeu 272

QY 1109 CATTTTGATGAAGCTGAGGATGTATTCTCTAATGAATTTTAAAGATCTCTAAATCCAGTT 1168
 DB 273 TyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrpHisAsnThrThr 292
 QY 1169 GTATATGAGAGTGTATTCAGCACTTCCAGTAACAATTTTCAAGGGATCAGCGTGTGTATGAT 1228
 DB 293 PhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAlaIleCysGluTyr 312
 QY 1229 AGCATAGTATGTAGAGAGGCTGTTCTTGGTTCATATGCCACAGGATGGACCCCAAC 1288
 DB 313 GlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyrHisGluGluAla 332
 QY 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAACCTTGTCCCAAGC 1348
 DB 333 GlnLysTrpAspArgTyrThrAspProValProSerProArgProGlySerCysIleAsn 352
 QY 1349 AAAACATTT-----GGTGGTTTGTACTCTACAAGGACCTTCTCTGATGTATTATAACC 1402
 DB 353 AsnTrpHisArgArgHisGlyTyrThrSerSerLeuGluLeuProAspAsnIleLeuAsn 372
 QY 1403 TTTCGAAGAAGTCAATCAGCCATGTACAATCCAGTGTCTCTATGAACAATCGCCCAATA 1462
 DB 373 PheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrpSerArgProLeu 392
 QY 1463 GTGATCAAAACGGATGTAAATTTATCAATTTTACACAAATTTGCTAGACCGAGTGGATGCA 1522
 DB 393 LeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAspArgValThrGly 410
 QY 1523 GAAGATCGA---CAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGCTTTTAAA 1579
 DB 411 LeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGlyTrpLeuLys 430
 QY 1580 GTAGTTTCAATCTCAAGGAGACTTGGTATGATTAGTAAGAGGTTCTGCTCGAAGAAATG 1639
 DB 431 AlaValSerLeu-----GlyProTrpValHis-----LeuIleGluGluLeu 444
 QY 1640 ACAGTTT---CGGGAACCGACTGCTATTTCACAATGGAGCTTTCACCTAAGCAGCAA 1696
 DB 445 GlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSerGlnSerLys 462
 QY 1697 CAATATATATGTTTCAACGGCTGGGTGGTCCCGAGCTCCCTTTACACCGGTGTGATATT 1756
 DB 463 LeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAlaAspCysMetLys 482
 QY 1757 TACGGGAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTGCTGGATGGT 1816
 DB 483 Tyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCysAlaTrpSerVal 501
 QY 1817 TCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCACAGACGA----- 1864
 DB 502 Asn-----ThrSerArgCysValAlaVal 509
 QY 1865 -----CAAGATATAAGAAATGGAGACCCCTGACTCAC 1897
 DB 510 GlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAspThrSerGlyLe 529
 QY 1898 TGTTCAGACTTACACCATGATAATCACCATGGCCACAGC-----CCTGAGAGAGAGA 1948
 DB 530 Cys-----AsnLeuArgGlySerLysValArgProThrProLys 543
 QY 1949 ATCATCTATGCTAGAGAAATAGTAGCACATTTTGGAAATGCAGTCCGAAGTCGCAGAGA 2008
 DB 544 AsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeu 563
 QY 2009 GCGTGTGCTTATTCGCAATTCACAGAGGCGAAAT-----GAAGAGCGGAAAGAGAGATC 2062
 DB 564 AlaHisAlaArgTrpThrPheGlyArgAspLeuProAlaGlnProGlySerPhe 583
 QY 2063 AGAGTGGATGATCATCATCAGCAGATCAAGGCTTCTGCTACGTAGTCTACAACAG 2122
 DB 584 LeuTyrAspAlaArgLeu-----GlnAlaLeuValValMetAlaAlaGlnPro 599

Qy 2123 AAGGATTGAGCAATTACCTCTGCCATGCGGTGGGAACATGGG 2164
::: :::: ||| ||| :::: |||
Db 600 ArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 613

Search completed: August 3, 2003, 10:21:13
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: August 3, 2003, 09:57:04 ; Search time 27 Seconds
(without alignments)
8490.379 Million cell u

Title: US-09-774-490-1

Perfect score:

Sequence: 1 aatcttttatttatcgatg.....aggcttttttcctaataacc 2709

Scoring table:

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 6574334

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

- MODELS=frame, nfp, model, - DEV=xlh
- Q/cgnt_1/USPTO_spool/US07774490/runat_03082003_095416_25355/app_query.fasta_1.2887
- DB=Issued Patents AA -QFMT=fastan -SUFFIX=xai -MINMATCH=0.1 -LOOPCL=0
- LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MAYRIX=blosum62 -TRANS=human40.cdi
- LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
- MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
- USER=US07774490.@CGN_1_30.@runat_03082003_095416_25355 -NCPU=6 -ICPU=3
- NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
- DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
- FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	4201	85.8	771	1	US-08-121-713D-54	Sequence 54, Appl
2	4201	85.8	771	1	US-08-835-268-54	Sequence 54, Appl
3	4201	85.8	771	2	US-09-060-692-54	Sequence 54, Appl
4	4201	85.8	771	3	US-08-833-391-54	Sequence 54, Appl
5	4201	85.8	771	4	US-09-060-610-54	Sequence 54, Appl
6	4201	85.8	771	5	PCR-US94-10151A-54	Sequence 54, Appl
7	3568	72.9	655	4	US-08-556-422A-3	Sequence 3, Appl
8	2562	52.3	477	1	US-08-136-922-2	Sequence 2, Appl
9	2030	41.5	775	4	US-09-308-179B-1	Sequence 1, Appl
10	982.5	20.1	862	4	US-08-556-422A-2	Sequence 2, Appl
11	838.5	17.1	607	4	US-08-556-422A-4	Sequence 4, Appl
12	785.5	16.0	724	1	US-08-121-713D-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-08-121-713D-54
; Sequence 54, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36, 627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEFAX:

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 771 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-121-713D-54

Alignment Scores:

Pred. No.: 0 Length: 771

Score: 4201.00 Matches: 771

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 85.80% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-121-713D-54 (1-771)

QY 200 ATGGCTGTTAACTAGGATTCTGCTTTCTGGGAGTATTACTTACAGCAAGAGCA 259

DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAGGCTGAAATATCTCAAAAGAAATGTG 319

DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATCTTC 379

DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439

DB 61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATCAAGATTTCCTCAAGATTGTGTGCCAGTATCTTACACCAAGA 499

DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTCGAAGATGTGCTAATTTTCATCAAGTA 559

DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATAATCAGACTCCTGTGACGCTGTGGAACGGGGCTTTTCATCCAATT 619

DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TCACCTACATTGAAATGGACATCATCTCGAGCAATATTTTAACTGGAGAACTCA 679

DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160

QY 680 CATTTTGAACGGCTGGGAAGAGTCCATATGACCCCTAAGCTCTGCACAGCATCCCTT 739

DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180

QY 740 TTAATAGATGGAAGATTACTCTGGAAGTGGAGTCTGATTTTATGGGGAGACTTTGCT 799

DB 181 LeuIleAspGlyLysLeuTyrSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200

QY 800 ATCTTCCGAACCTTTGGGCACCAACCAATCAGGACAGACAGCATGATTCAGAGTGG 859

DB 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 860 CTCATGATCCAAAGTTCTATAGTCCACCTCATCTCAGAGAGTGACATCTCCTGAAGAT 919

DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTATCTTTTCTCGTGAATAATGCAATAGATGGAGAACACTCTCGAAAGCT 979

DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260

QY 980 ACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCAGCAAGTCTGGTG 1039

DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAATGGACACATTCTCTCAAGCTGCTGATTTGCTCAGTCCAGGTCCTCAATGCG 1099

DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCATTTTGTATGAACACTGCAGGATGATTTCTTAATGAATCTTAAAGATCCTAAA 1159

DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320

QY 1160 AATCCAGTTGATATGAGTGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219

DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGATATGATGATGATGAGAGGCTGTTCTTGTGCTCATATGCCACACAGGAT 1279

DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCCCAACTATCAATGGTGCCTTTATCAAGAAAGAGTCCCTATCCACGCCAGGAAT 1339

DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCGACGCAAAATTTGGTGGTTTGGACTCTCAAGAGGACCTTCTGTATGATGTTATA 1399

DB 381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400

QY 1400 ACCTTTGCAAGAACTCATCCAGCCATGTACAATCCAGTGTCTTCTATGAAACAATCGCCCA 1459

DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420

QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519

DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

QY 1520 GCAGAGATGACAGATGATGTTATGTTTATCCGACAGATGTGGACCGTCTCTTAAA 1579

DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639

DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480

QY 1640 ACAGTTTTTCCGGAAACCGACTGTATTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACAA 1699

DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATGTTTCAAAGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759

DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520

QY 1760 GGGAAAGCTGCTGCTGAGTGTGCTGCCGAGACCCCTTACTGTGCTTGGGATGGTTCT 1819

DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540

QY 1820 GCATGTTCTGCTATTTTCCCACTGCCAAGAGACGCAACAGCAGCAAGATATAAGAAAT 1879

DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560

QY 1880 GGAGACCCACTGACTCTGTTTACAGATTACCATGATATCACCATTGCGCCACAGCCCT 1939

DB 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580

QY 1940 GAACAGAGAAATCATCTATGTTGAGTAATAGTAGACATTTTGGAAATGCAGTCCGGAAG 1999

DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600

QY 2000 TCGCAGAGAGCGCTGCTCTATTGGCAATTCAGAGGCGGAAATGAAGACGCAAGAAAGAG 2059

DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620

QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTAGCTAGTCTACA 2119

DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640

QY 2120 CAGAGGATTCAGGCAATTAACCTCTGCATCGGTGGACATGGGTTTCATCAAACTCTT 2179
Db GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCAATTTGGAGAACTCTTTCATAAAGAT 2239
Db LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG 2299
Db AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTCACAGACTTCATGACAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTCGAACAGTTTGGAAAAGGACCGAAACAACTCGCAAGGCCAGGACATACC 2419
Db PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CAGGGAACAGTAACAATGGAGCACTTACAGAAATAAAGAGGTAGAAACAGAGG 2479
Db ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACGATTTGAGAGGCCACCCAGGAGTGTC 2512
Db ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

US-08-835-268-54
; Sequence 54, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-835-268-54

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-835-268-54 (1-771)

QY 200 ATGGGCTGGTTAACTAGAGTTGTCTGCTTTTGGGGAGTATTACTTACAGCAAGAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGAACAAATGTCACAGGCTGAATATCTCTCAAAAGAAATGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTGTGTGGCCAGTATCTTACACCAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGGCTGGAAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTTGTACGCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGACATCATCTTGAGGACAAATATTTTAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTCAAAAACGGCCGTGGGAAGAGTCCATATCACCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATTTATCTCTGGAACTCCAGCTCAATTTTATGGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACTCTTGGGCACCAACCCCAATCAGGACAGACGACATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATGATCAAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGAGAACTCTGGAAGAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACGCTAGATAGTCAAGATATGCAAGAAATGACTTTTGGAGGGGACAGAACTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTGATGAACCTGCAGGATGTATTCTTAATGAACCTTTAAAGATCTCTAAA 1159

Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLeuAspProlys 320
QY 1160 AATCAGTTGTATATGAGTGTATACAGTCTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLeuGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGATGAGAGGGTGTCTTGGTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAAC 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGAAACATTTGGTGGTTTGAATCTACAAAGACCTCTTCGTATGATGATTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCNAGAGTCCAGCCATGTACATCCAGTGTCTTCCCTATGAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTAATCAATTTACACAAATTTGCTAGACCGAGTGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGGAAGATGATGATGATTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCATTTCTTAAGAGACTGCTGATGATTTAGAGAGTTCTCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTTCGGAAACCGACTGCTATTTTCCAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGGTTCAACGCTGGGTTCGCCAGTCCCTTTTACACGGTGTGATATTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTACTGCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCACAAAGACGACAAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGTTTCAAGATTTACACATGATATACATGACATGCGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGATGCGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProlys 600
QY 2000 TCGCAGAGACGCTGCTATTGGCAATTTCCAGAGCGCAATCAAGAGCAAAAGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrPheGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGCTAGTGTATACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATTTACCTCTCCATGCGGTGGAACATGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTACCTCGAAGTCTATTGACACAGAGCATTTTGGAAAGAACTTCTTCATAAGAT 2239

Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATCGAGATGCTCTTAAGACCAAAATGTCCTCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATCGAGCTCATCAACACACCCCAATCTCAACACGATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTCAACAAGTTTCGAAAAAGGACCGAAAAACAACGTGCGCAAGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTACAAATGGAACCACTTCAACAAAAATAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAACAATTTGAGAGCGCACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3
US-09-060-692-54
; Sequence 54, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-692-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 2 Gaps: 0
 US-09-774-490-1 (1-2709) x US-09-060-692-54 (1-771)

QY 200 ATGGCTGGTTAACTAGGATTTCTGTCTTTCTGGGAGATTTACTTACAGCAAGAGA 259
 DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGCGCTGAAATATCTCAAGAGAAATGTTG 319
 DB 21 AsnTyrGlnAsnGlyValAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCAAACAATGTGATCACTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTTC 379
 DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
 DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTGTCGCCAGTATCTTACACAGAGA 499
 DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArg 100
 QY 500 GATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGCTTAATTTTCATCAAGGTA 559
 DB 101 AspGluCysLeuTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAAAGGGGGCTTTTCATCCAAT 619
 DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTGAAATTTGGACATCATCTCGAGGACAAATATTTTAAAGCTGGAGACTCA 679
 DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAAACGGCGTGGGAAGAGTCCATATGACCCCTAAAGCTCTGACAGCATCCCTT 739
 DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTAATCTCTGGAACTCGAGCTGATTTTATGGGCGAGACTTTGCT 799
 DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAATCTTTGGCCACCACCACCAATCAGGACAGACGATGATTCAGGTGG 859
 DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCATGATCCAAAGTTTCATTAGTCCCACTCTCATCTCAGAGAGTGACAAATCCTGAAGAT 919
 DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATCTTTTCTCGTGAATAATGCAATAGATGAGAACACTCTGGAAAAGCT 979
 DB 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTACGCTAGATAGTTCAGATGCAAGATGACTTTGGAGGGCCACAGAGTCTGGTG 1039
 DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACAATCTCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCCTCAATGGC 1099
 DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTTGATGAATCTGCAGGATGTATTCCTTAATGAATTTTAAAGATCCTTAA 1159
 DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGTATGGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATCAGCGGTG 1219
 DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGTATAGCATCAGTGATGTGAGAAGGTGTTCTTGTGTCATATGCCACAGGGAT 1279
 DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCAACTATCAATGGTGCTCTTATCAAGAGAGTCCCTATCCAGGCCAGAACT 1339
 DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCACAGAAAACATTTGGTGGTTTTCACCTACAAAGGACCTTCTCTGATGATGTTATA 1399
 DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGCCAAGAAGTCATCCAGCCATGTACCAATCCAGTGTTCCTATGAACAATCGCCA 1459
 DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAAACGGATGTAAATTCATATTTACAAAATTCGTGTAGACCGAGTGAT 1519
 DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGCACAGTATGATGTTATGTTATCGAACAGATGTTGGACCGTCTCTTAAA 1579
 DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTCTCGAAGAAATG 1639
 DB 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAAACCGACTGTATTTTACCAATGGAGCTTTCCACTAAGCAGACAACAA 1699
 DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGATATTTAC 1759
 DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTGCTTGGGATGGTTCT 1819
 DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAAGACGACAGATATAGAAT 1879
 DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
 QY 1880 GGACACCCACTGACTCTGTTACAGCTTACACCATGATATCACCATGGCCAGACCCCT 1939
 DB 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAATCATCTATGTTGTAGAGAATAGTAGCACATTTTGGAAATGCGATCCGAG 1999
 DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAATGAAGACGCAAGAAAGAG 2059
 DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
 QY 2060 ATCAGCTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTACGTAGTCTACAA 2119
 DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGGATTCAGGCAATTTACCTCTGCGATGCGGTGGAACATGGGTTTCATACAACTCTT 2179
 DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGTAAACCTGGAGGTCATTGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 2239
 DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
 QY 2240 GATGATCGAGATGGCTCTAGACCAAGAAATGTCCAATAGCATGACACCTAGCAGAG 2299
 DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700

QY 2300 GTCTGGTACAGAGACTTCATGAGCTCATCAACCCCACTCAACACGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAGGGACCGAAACCAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAACAATGGAGGACCTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCCACCCAGGAGTGTG 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 4

US-08-833-391-54
; Sequence 54, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthews, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-391-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 3 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-833-391-54 (1-771)

QY 200 ATGGCGTGTAACTAGGATTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGAGA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAAGAACAAATGTCCAAAGGCTGAAATTTATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACATGTATCATCTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCCCTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGGAACGAGTAGCTGTATCTTCGAGCAAAAGGATCAATATTTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATCAATGCAAGTGGGCTGGNAAGACATCTCTGAAGAAATGTGCTAATTTTTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTTCATCCAATT 619
Db 121 LeuLysAlaLysAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProlle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAATATTTTAAAGCTCGAAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTGAAAACGGCCGTGGAGAGCTCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACACTGATTTTATGGGCGAGACTTTCCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGGCGACCAACCCCAATCAGGACAGACAGCATGATTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATAGTGGCCACCTCATCTCAGAGAGTGACAACTCCGAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLysSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAATAGGTACAGATATGCAAGAAATGACTTTTGGAGGCGACAGAAAGTCTGGT 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCCAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTAGTAAGCTCAGGATGTATTCCTTAATGAACCTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATATTTTCAAGGGATCAGCGGT 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGATGTATAGCATGATGTGAGAAAGGTGTTCCTTCGTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGGTGCCTTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGAAC 1339

Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCACGACAAACATTTGGTGGTTTGAATCTACAAAGAGCCTTCCCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGACAAATCCAGTGTTCCTATGAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGATTAATATTCAATTTACAAATTTGCTAGACCGAGTGCAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGACACATGATGATTTATCGGAACACAGATGTTGGGACCGCTTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGACCGACTCTATTTACGAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACCGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGAAGCGTGTGAGTGTGCTCCCGAGACCTTACTGTGCTGGATGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
QY 1820 GCATGCTTCGCTATTTCCCACTCCAAAGAGACCCACAGACGACAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTGCTGTTACAGCTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGTTAGAGATAGTAGCACATTTTGGATGCGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCGCTGCTGTTGCGCAATTCAGAGCGCAATGAAGACGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAGTCTCAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGCAATTAATCTCTGCTGCGTGGAAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTATTGACACAGAGCATTTGGAAGAACTTCTTCATTAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGACTTCATGCGCTCATCAACACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAGGACCGAAACAAACGTCGGCAAGGCCAGGACATACC 2419

Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGAAGCAGCTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACAGATTTGAGAGGCGCACCCAGGAGTGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 5
US-09-060-610-54
; Sequence 54, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/060.610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-610-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-060-610-54 (1-771)

QY 200 ATGGGCTGGTTAACTAGCATTTGCTCTTTCTGGGGAGTATTACTTACAGCAAGAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGAAACAATGTGCCAAGGCTGAAATTTATCTCAAGAAATGTTG 319

Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAAGTGTATCACTTTCAATGGCTTGGCCACAGCTCCAGTATATACACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGAAACGAGTAGGTGTATGTGGAGCAAGGATCACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAAATATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArg 100
QY 500 GATGAATGCAAGTGGCTGGAAGACATCCTGGAAGATGTCTAAATTTTCATCAAGTGA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAATTCGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTGTGAACGGCGTGGAGAGTCCATATGACCTTAAGCTGCTGACAGCCTCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAGAATTATCTGGAACCTGACGCTGATTTTATGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTCTTGGGCACCAACCAATCAGACAGACAGCAGATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATGATCCAAAGTTCAATTAGTCCCACTTCATCTCAGAGAGTGACAACTCCTGAAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ATCAGCTAGAAATAGTGCAGATATGCAAGATGACTTTGGAGGGCAGCAAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAATGACACATCTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAATGGC 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGATCAACTGCAAGTATGTTCTTAATGAACTTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTATGACTTCCAGACTTCCAGTAACATTTTCAAGGATCAGCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGATGAGAGGTGTTCTGTGTCATATGCCCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTCTACAAGAGACCTTCTCGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400

QY 1400 ACCTTTGCAAGAAAGTCATCCAGCCATGTACAATCCAGTGTTCCTTATGAAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATACTGATCAAAACCGATGTAAATATCAATTTTACAAATTTGCTGCTAGACCGAGTGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGACAGTATGATGTTATGTTTATCGACACAGATGTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATCATTTTAGAAGAGGTCTCTCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTTCGGGAACCGACTGTATTTTCACAAATGGAGCTTTTCCATTAAGCAGACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGGTTCAACGGCTGGGTGCGCCAGCTCCCTTTTACACCGTGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTCTGAGTGTGCTCGCCCGACACCTTACTGTGCTTGGATGGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAAGACGCAAGATATTAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTACTGTTACACTTACACCATGATATCACCATGGCCAGACCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAACAGAGAATCATCTATGTTGTAGATAATAGTAGACATTTTGGAAATGACGATCCCGAAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGGTCTATTGGCAATTCAGAGGCGAATGAAGACGCAAGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGSCCTTCTGCTACGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACTCTGCTCATCGGTGGAACATGGGTTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAGTCAATTGACACAGAGCAATTTGGAAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATGGAGATGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATGCTCATCAACCAACCCCAATCTCAACACATGGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGAGCCGAAAAACAACGTTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGACAGTAACAATGGAACCACTTACAAGAAATAGAAGGTAGAAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760

QY 2480 ACCCAGCAATTTGAGAGGCGACCCAGAGTGTTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 6

PCT-US94-10151A-54
; Sequence 54, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FLEHR HOBBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10151A
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-10151A-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US94-10151A-54 (1-771)

QY 200 ATGGCGTGTAACTAGAGTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGCAAAATGTGCAAGGCTGAAATTTATCTCAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluWetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGACCGAGTAGGCTGTATGTTGGAGCAAGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 CACCTGGTTAATATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATGCAAGTGGGCTGGAAGAAAGACATCTCTGAAGAAATGTGCTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATACAGACTCAGTGTGACGCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAACTCGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACGGCGCTGGGAAGAGTCCATATAGCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGGAGACTTTCCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGGCACCCACCATCAGGACAGGACAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTCATTTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGCGTGAANAATCAATAGATGAGACACTCTCGAAGACT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTGTGAGATATGCAAGAAATGACATTTTGGAGGCGACAGAAAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAAATCTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGCG 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTAGTACGAGATGATTTCTTATGAACTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGACTGATGTGAGAGGTTGTTCTTGTTCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTCATCAATGGCTGCCTTATCAAGGAGAGTCCCTATCCACGCCAGGAACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTGTACTCTCAAGGACCTTCTCTGATGATTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCCCATGACAATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGAGTGAATTTATCAATTTTACAAATTTGTCGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATCGACAGTATGATGTTATGTTTATCGAACACAGATGTTGGAGCGCTTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTTAGAAGAGGTTCTGCTCGAAGAAATG 1639

```
|||||ValValSerIleProLysGluThrTrpTyAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACATGAGCAGCAAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGGTTCAACGCTCGGGTCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
Db 501 LeuTyriIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GCGAAAGCGTGTCTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGGATGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGlySer 540
QY 1820 GATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGACAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACCTGCTTACAGCTTACACCTTACACCTGATTAATCACCATGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAATCATCTATCGTGTAGAGATAGTACACATTTTTCGAATGCGATCCGAAG 1999
Db 581 GluGluArgIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TGCAGAGACGCTGCTGCTATTGGCAATTCAGAGCGCAATGAAGAGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyTrpGlnPheGlnArgArgAsnGluGluArgLysGlu 620
QY 2060 ATCAGAGTCGATCATCATATCAGCAGACATCAAGGCTTCTGCTAGTGTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTACAGGAATTTACCTCTGCCATGCGGTGGAACATCGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTGGAAGTCAATTGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspCysLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGACAGACTTCTATGCGCTCATCAACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTrpTyArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAGGGCCGAAACACAGCTGCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysCysLysArgAsnArg 760
QY 2480 ACCCAGAAATTTGAGAGGCGCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
```

RESULT 7

```
US-08-556-422A-3
; Sequence 3, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
```

```
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556.422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-556-422A-3

Alignment Scores:
Pred. No.: 0 Length: 655
Score: 3568.00 Matches: 655
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.88% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-3 (1-655)
QY 200 ATGGCGTGTAACTAGAGATTGCTGCTTTCTGGGAGATATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAACAATGTCACAGGCTGAAATATCTCAAAAGAAATGTTG 319
Db 21 AsnTyGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyLysGluMetLeu 40
QY 320 GAATCCAAACATGTGATCATTTCATCAATGGCTTGGCCCAACAGCTCCAGTTATCATCCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyHisThrPhe 60
QY 380 CTTTGGATGAGAAACGAGTAGCTGTATGTTGGAGCAAAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACTGTGTTAATATCAAGGATTTTCAAAGATTCTGTGGCCAGTATCTTACACCAAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyTrpArgArg 100
QY 500 GATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCATTGTACCCCTGTGGAACGGGGCTTTTCATCCCAATT 619
Db 121 LeuLysAlaTyAsnGlnThrHisLeuTyAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAGCTCGAGAACTCA 679
Db 141 CysThrTyriIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyArgProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTCAGACTGATTTTATGGGGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTySerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTGGGACCCACCACCATCAGGACAGACAGATGATTCAGAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisPheIleProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGATATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAGCT 979
Db 241 AspLysValTyPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
```

QY 980 ACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTTGGAGGCGCACAGAAGTCTGGTG 1039
Db ThrHisAlaArgIleGlyGlnIleCysLysAsnAepPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGCAACAATCTCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCCTCAATGGC 1039
Db AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTGTAAGTCACTGAGGATGATTCTTAATGAACCTTTAAAGATCTCAAA 1159
Db IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProlys 320
QY 1160 AATCAGTTGTATGAGTGTTCAGACTCCAGTAAACATTTCAAGGATCAGCGGTG 1219
Db AsnProValValTyGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGATGAGAGGGTGTCTTCTGTGTCATATGCCACAGGGAT 1279
Db CysMetTySerMetSerAspValArgValPheLeuGlyProTyAlaHisArgasp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAT 1339
Db GlyProAsnTyGlnTrpValProTyGlnGlyArgValProTyProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGAATCTCTCAAGAGACCTTCTGTGATGTTTATA 1399
Db CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAGAACTCATCAGGCATGACCAATCCAGTGTTCCTATGAACAATCGGCCA 1459
Db ThrPheAlaArgSerHisProAlaMetTyAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGGATGTAATATCAATTTACAAATTTGCTAGACCGAGTGAT 1519
Db IleValIleLysThrAspValAsnTyGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGTGCAGATGATGTTTATGTTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db AlaGluAspGlyGlnTyAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATCTCAGGAGACTGTGATGATTAATAGAGAGCTTCTGCTGGAGAATG 1639
Db ValValSerIleProLysGluThrTrpTyAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGTATTTACGAATGAGCTTCCACTAAGCAGCAACA 1699
Db ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGTGGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
Db LeuTyIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GGGAAAGCTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGATGGTGTCT 1819
Db GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCCTATTTCCCACTGCAAGAGAGCGCAAGACGACAGATATAGAAT 1879
Db AlaCysSerArgTyPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTCTGTTGAGACTTACCATGATATCATCCATGGCCAGCGCCT 1939
Db GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGTTGATAGATAGTACACATTTTGGATGCAGTCCGAAG 1999
Db GluGluArgIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProlys 600
QY 2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGAATGAAGACGGAAGAAGAG 2059
Db SerGlnArgAlaLeuValTyTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620

QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGCCTTCTGCTAGTACTTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGTTCAGCAATTAACCTCTGCATGCGGTGGACATGGG 2164
Db 641 GlnLysAspSerGlyAsenTyLeuCysHisAlaValGluHisGly 655
RESULT 8
US-08-136-922-2
; Sequence 2, Application US/08136922
; Patent No. 5416197
; GENERAL INFORMATION:
; APPLICANT: Raper, Jonathan A.
; APPLICANT: Luo, Yuling
; TITLE OF INVENTION: Compositions Which Regulate Neural
; TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: NO. 541619718
; STREET: One Liberty Place
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,922
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-1428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-136-922-2
Alignment Scores:
Pred. No.: 1,01e-246 Length: 477
Score: 2562.00 Matches: 472
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 3
Query Match: 52.33% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-136-922-2 (1-477)
QY 899 GAGATGACAAATCCTGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGAT 958
Db 2 GluHisAspAsnProGluAspAspLysValTyPhePheArgGluAsnAlaIleAsp 21
QY 959 GGAGACACTCTGGAAAGCTACTCAGCTAGATAGGTGATATGCAAGATGCAAGATGACTTT 1018
Db 22 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 41
QY 1019 GGAGGGCACAAGATCTGGTGAATTAATGACAACTTCTCAAAGCTCGTCTGATTGTC 1078
Db 42 GlyGlyHisArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCys 61
QY 1079 TCAGTGCCAGGTCCTCAAAATGCGATTCATCTATTTTATGAACTGCAGGATGATTCTCTA 1138

Db 147 ValGlyHisHisSerGluGluProLeuPhePheHisLeuGluSerHisArgSerGluArgGly 166
QY 695 COTGGGAAGAGTCCATATACACCTTAAGCTGCTACAGACATCCCTTTTAAATAGATGAGAA 754
Db 167 ArgGlyArgCysProPheAspProAsnSerSerPheValSerThrLeuValGlyAsnGlu 186
QY 755 TTATACTCTGGNACTGCGAGCTGATTTTATGGGCGAGACTTGTCTATCTTCGGAACCTTT 814
Db 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpGlyArgAspSerAlaIlePheArgSerMet 206
QY 815 GGGCACCACCCACCAATCAGGACAGCAGCATGATCCAGGTGCTCAATGATCCAAAG 874
Db 207 GlyLysLeuGlyHisIleArgThrGluHisAspAspGluArgLeuLeuGluLeuProlys 226
QY 875 TTCATTAGTGCCACCTCATCTACAGAGTACAACTCTGGAAGTACGACAAAGTATATCTTT 934
Db 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysMetTyrPhe 246
QY 935 TTCTTCCGTGAATGCAATAGATAGAGACACACTCTGGAAGAAGTACTACAGCTAGATA 994
Db 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAsnAlaHisThrIleTyrThrArgVal 266
QY 995 GGTCAATATGCAAGAATGACTTTTCGAGGCGACAGAGTCTGCTGAATAAATGACACAA 1054
Db 267 GlyArgLeuCysValAsnAspMetGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
QY 1055 TTCTCAAAGTCTGCTGATTTGCTAGTCCAGTCCAAATGCGATTGACACTCATTTT 1114
Db 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306
QY 1115 GATGACTGCGAGGATGATCTTCTTAATGAATTTAAGATCTCTAAATATCCAGTTGTATAT 1174
Db 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspProLysAsnProValIlePhe 326
QY 1175 GGAGTGTTCAGACTTCCAGTAACATTTTCAAGGGATCAGCGCTGTGATGTATAGATG 1234
Db 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaValCysValTyrHisMet 346
QY 1235 AGTGATGTGAGAAGGTGTTCTTGGTCCATATGCCAGGATGAGACCCAACTATCAA 1294
Db 347 SerSerIleArgGluAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
QY 1295 TGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCGCAGAACTTGTCCGCAAGAACAA 1354
Db 367 TrpSerLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
QY 1355 TTTGTGTGT---TTTGACTCTACAAAGACCTTCTGATGATGTATTAACCTTTGCCAGA 1411
Db 387 AsnGlyGlyLysTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
QY 1412 AGTCATCCAGCCATGTATCAATCCAGTGTTCCTATGAACAATCGCCCAATAGTATCAA 1471
Db 407 MethHisProLeuMetTyrGlnProIleLysProValHisLysLysProIleLeuValLys 426
QY 1472 ACGGATGTAATATCAATTTACAAATTTGCTGAGACCGAGTGGATGCGAGAGATGGA 1531
Db 427 ThrAspGlyLysTyrAsnLeuArgGlnLeuAlaValAspArgValGluAlaGluAspGly 446
QY 1532 CAGTATGATGTATTTTATCGGAACAGACTTGTGGACCGTCTTAAAGTAGTTTCATTT 1591
Db 447 GlnTyrAspValLeuPheIleGlyThrAspThrGlyIleValLeuLysValIleThrIle 466
QY 1592 CCTAAGGAGACTTGGTATGATTTAAGAGAGTCTTCTGGAAGAAGTACACAGTTTTCGG 1651
Db 467 TyrAsnGlnGluThrGluTyrMetGluGluValIleLeuGluLeuGluGlnIlePheLys 486
QY 1652 GAACGAGCTGCTATTTACGAATGAGGTTTCCACTAAGCAGCAACAACTATATATTTGGT 1711
Db 487 AspProAlaProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506
QY 1712 TCACGGCTGGGTGGCTCCCTTACCGGTGTGTATATTACGGGAAGACGTGT 1771
Db 507 SerAlaSerAlaValAlaGlnValArgPheHisCysAspMetTyrGlySerAlaCys 526

QY 1772 GCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGCTGGGATGGTTCTGCATGTTCTCGC 1831
Db 527 AlaAspCysCysLeuAlaAspArgProTyrCysAlaTrpAspGlyIleSerCysSerArg 546
QY 1832 TATTTTCCACT-----GCAAAGAGACGACAGACGACAGACATATAGAAATGGA 1882
Db 547 TyrTyrProThrGlyAlaHisAlaLysArgArgPheArgGlnAspValArgHisGly 566
QY 1883 GACCACCTGACTCTGTTACAGCTTACACCATGATAATCACCATTGGCCACAGCCCTCAA 1942
Db 567 AsnAlaAlaGlnGlnCysPheGlyGlnPheValGlyAspAlaLeuAspArgThrGlu 586
QY 1943 GAGAAATCATCTATGTTGTTAGAGATAGTAGACATTTTGGAAATCCAGTCCGAAGTCG 2002
Db 587 GluArgLeuAlaTyrGlyIleGluSerAsnSerThrLeuLeuGluCysThrProArgSer 606
QY 2003 CAGAGAGCGTGTCTTATTTGGCAATTCAGAGGCGAAATGAAGAGCGGAAAGAGAGATC 2062
Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgAspValArgLysGluGluVal 626
QY 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAACAG 2122
Db 627 LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgValArgLys 646
QY 2123 AAGGATTACGCAATTTACCTCTGCTCGGTGGAACATGGGTTTCATACAACTCTTCTT 2182
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisAsnPheValHisThrValArg 666
QY 2183 AAGTAACCTTGGAAAGTCAATTCAGACAGACATTTGGAAGAACTTCTTCATAAAGATGAT 2242
Db 667 LysIleThrLeuGluValValGluGluHisLysValGluGlyMetPheHisLysAspHis 686
QY 2243 GATGAGATGCTCTTAAGACCAAGAAATGTCC-----AATAGCATGACACCT 2290
Db 687 GluGluGlu-----ArgHisHisLysMetProCysProLeuSerGlyMetSerGln 704
QY 2291 AGCCAGAAGTCTCGTACAGAGACTTTCATGAGCTCATCACACCCCAATCTCAACACG 2350
Db 705 GlyThrLysProTyrTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArg 724
QY 2351 ATGATGATGTTCTGTGAACAGTTTGGAAAGGACCGAAGAAACACGTCGCGCAAGGCCA 2410
Db 725 ValGluGluTyrCysGluLysValTyrCysThrAspLysLysArgLysLysLeuLysMet 744
QY 2411 GGACATATCCCGAGGAACAGTAACAATGGAAGCACTTACAGAAATAAGAAAGGTAGA 2470
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLys---Arg 758
QY 2471 AACAGAGGAGCCACGAAATTTGAGAGGCGCACCCAGG 2506
Db 759 LeuArgSerLysAlaGluHisPheArgLeuProArg 770

RESULT 10

US-08-556-422A-2
; Sequence 2, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DPN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Homo sapiens

Db 623 GlnCysLeuSerGluGluArgValLysAsnLysThrValPheGlnValValAlaLysHis 642
Qy 2189 -----ACCTCGAAGTCATTGAC 2205
Db 643 ValLeuGluValLysValProLysProValValAlaProThrLeuSerValValGln 662
Qy 2207 ACAGAG 2212
Db 663 ThrGlu 664

RESULT 11
US-08-556-422A-4
; Sequence 4, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-556-422A-4

Alignment Scores:
Pred. No.: 9,7e-75 Length: 607
Score: 838.50 Matches: 210
Percent Similarity: 49.84% Conservative: 106
Best Local Similarity: 33.12% Mismatches: 249
Query Match: 17.13% Indels: 69
DB: 4 Gaps: 22

US-09-774-490-1 (1-2709) x US-08-556-422A-4 (1-607)

Qy 356 AACAGCTCCAGTTATCATACCTCTCTTTGGATGAGAACGGAGTAGGCTGTATGTGA 415
Db 12 AsnLeuSerAsnTyrThrAlaLeuLeuSerGlnAspGlyLysThrLeuTyrValGly 31
Qy 416 GCAAAGGATACATATTTTTCATTCGACCTGTTAATATCAAG-----GAT 460
Db 32 AlaArgGluAlaLeuPheAlaLeuAsn----SerAsnLeuSerPheLeuProGlyGlyGlu 50
Qy 461 TTTCAAAAGATGTGTGGCCAGTATCTTACACCCAGAGAGATGAATGCAAGTGGGCTGGA 520
Db 51 TyrGlnGluLeuLeuTyrSerAlaAspAlaAspArgLysGlnGlnCysSerPheLysGly 70
Qy 521 AAAGACATCTCGAAGAATGTCTTAATTCATCAAGTACTTAAGGCGATATAATACAGACT 580
Db 71 LysAspProLysArgAspCysGlnAsnTyrIleLysLeuLeuProLeuAsnSerSer 90
Qy 581 CACTGTGTACGCTGTGGAGCGGGCTTTTCATCCATTTTCACCTACATTAATGAATGGA 640
Db 91 HisLeuLeuThrCysGlyThrAlaAlaPheSerProLeuCysAlaTyrIleHisIleAla 110
Qy 641 CATCATCTGAGGACAAATATTTTAAAGCTGGAG-----AACTCACATTTT 685
Db 111 Ser-----PheThrLeuAlaGlnAspGluAlaGlyAsnValIleLeu 124
Qy 686 GAAACGGCCCTGGGAAGATGCTATATGACCTTAAGCTGTGACAGACATCCCTTTTAATA 745
Db 125 GluAspGlyLysGlyHisCysProPheAspProAsnPheLysSerThrAlaLeuValVal 144
Qy 746 GATGAGATATTAATCTCTGGAACCTGCAGCTGATTTTATGGCGGAGACTTGTCTATCTTC 805
Db 145 AspGlyGluLeuTyrThrGlyThrValSerSerPheGlnGlyAsnAspProAlaIleSer 164

Qy 806 CGAACTCTTGGGCACCACCAATCAGGACAGACAGCATGATTCAGGTGGCTCAAT 865
Db 165 ArgSerGlnSer---SerArgProThrLysThrGluSer---SerLeuAsnTrpLeuGln 182
Qy 866 GATCAAAAGTTCATTAGTGCCCACTCATCTCAGAGAGT---GACAATCCT---GAA 916
Db 183 AspProAlaPheValAlaSerAlaThrSerProGluSerLeuGlySerProIleGlyAsp 202
Qy 917 GATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAAACACTCTGGA 976
Db 203 AsnAspLysIleTyrPhePheSerGluThrGlnGluPheGluPheGluAsn 222
Qy 977 GCTACTCACGCTAGATAGTGCATATGCAAGATGACTTTTGGAGGSCACAGAGTCTG 1036
Db 223 ThrIleValSerArgValAlaArgValCysLysGlyAspGluGlyGlyGluArgValLeu 242
Qy 1037 GTGAATAATGGACAACTTCCTCAAGCTCGTCTGATTCCTCAGTCCAGGTCCAAT 1096
Db 243 GlnGlnArgTrpThrSerPheLeuLysAlaGlnLeuLeuCysSerArgPro---AspAsp 261
Qy 1097 GGCATTGACACTCATTTTGTGATGAACCTGAGATGCTATTCTTAATG-----AACTTTAAA 1150
Db 262 GlyPhePro-----PheAsnValLeuGlnAspValPheThrLeuAsnProAsnProGln 279
Qy 1151 GATCCTAAATCCAGTTGTATATGAGTGTGTTTACGACTTCC-----AGTAACATTTTC 1204
Db 280 AspTrpArgLysThrLeuSerIleGlyValPheThrSerGlnTrpHisArgGlyThrThr 299
Qy 1205 AAGGATCAGCCGTGTGTATGTATAGCATGATGATGAGAGGGTTCCTTGGTCCA 1264
Db 300 GluGlySerAlaIleCysValPheThrMetAsnAspValGlnLysAlaPheAspGlyLeu 319
Qy 1265 TATGCCACACGGGATGGACCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTAT 1324
Db 320 TyrLysLysValAsnArgGluThrGlnGlnTrpTyrThrGluThrHisGlnValProThr 339
Qy 1325 CCACGGCCAGGAACCTTGT-----CCAGCAAAAACATTTGTGTGTTTTGACTCTACAAG 1378
Db 340 ProArgProGlyAlaCysIleThrAsnSerAlaArgGluArgLysIleAsnSerSerLeu 359
Qy 1379 GACCTTCCTGATGATGTTTAACTTTTGCAGAAAGTATCATCCAGCATGTACATCCAGTG 1438
Db 360 GlnLeuProAspArgValLeuAsnPheLysLysAspHisPheLeuMetAspGlyGlnVal 379
Qy 1439 TTTCTATGAACATCCCAATAGTATCAAAACGGATGTAAATTTCAATTTACACAA 1498
Db 380 -----ArgSerArgLeuLeuLeuGlnProArgAlaArgTyrGln-----Arg 394
Qy 1499 ATTGCTCTAGACCCAGTGGATGCAGAAAGATGCAGATATGATGTTATGTTTATCGGAACA 1558
Db 395 ValAlaValHisArgValProGlyLeuHisSerThrTyrAspValLeuPheLeuGlyThr 414
Qy 1559 GATGTGGGACCGTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAA 1618
Db 415 GlyAspGlyArgLeuHisLysAlaValThrLeuSerSerArgValHis-----430
Qy 1619 GAGTTCTGTGGGAAGAAATGACAGATTTTTCGGAACCGACTGCTATTTACGCAATGGAG 1678
Db 431 -----IleIleGluGluLeuGlnIlePheProGlnGlnProValGlnAsnLeuLeu 448
Qy 1679 CTTTCCACTAAGCAGCAACAACTATATATTGTTTCAACGCTGGGGTGTGCCAGCTCCCT 1738
Db 449 LeuAspSerHisGlyGlyLeuLeuTyrAlaSerSerHisSerGlyValValGlnValPro 468
Qy 1739 TTACACCGGTGTGATATTTACGGGAAGCGGTGCTGAGTGTGCTCGCCCGAGACCT 1798
Db 469 ValAlaAsnCysSerLeuTyr---ProThrCysGlyAspCysLeuLeuAlaArgAspPro 487
Qy 1799 TACTGTCTTGGATGGTCTGTCATGT-----TCTCGCTATTTTCCCACTGCAAG 1849
Db 488 TyrCysAlaTrpThrGlySerAlaCysArgLeuAlaSerLeuTyrGlnProAspLeuAla 507

Qy	1073	ATTGTGCTCAGTGCAGGTC	CAAAATGGCATTGACACTCATTTGATGAACATGCAGGATGTA	1133
Db	314	AsnCysSerIleSerGlyGlu-----PheProPheTyrPheAsnGluIleGlnSerVal	331	
Qy	1133	TTCCTAATGAACATTTAAAGATCCT	AAAAATCCAGCTGTGTATATAGCATGTGATGGAGTGTTTACGACTTCC	1192
Db	332	TyrGlnLeuProSerAspLysSerArg-----PhePheLaThrPheThrThrSer	348	
Qy	1193	AGTAACACATTTTCAAGGGATCAGCCGCTGTGTATATAGCATGTGATGGAGTGTGAGAGAGGGTG	1252	
Db	349	ThrAsnGlyLeuIleGlySerAlaValCysSerPheHisIleAsnGluIleGlnAlaAa	368	
Qy	1253	TTCCTTGGTCATATGCCACAGGAGTGGACCCAACTATCAATGGGTGCTC	1309	
Db	369	PheAsnGlyLysPheLysGluGlnSerSerSerAsnSerAlaTrpLeuProValLeuAsn	388	
Qy	1310	GGAGAGTGCCTTATCCACGCGCAGGAACTGTGCCAGCAAAACATTTGGTGGTTTGAC	1369	
Db	389	SerArgValProGluProArgProGlyThrCys-----ValAsn	401	
Qy	1370	TCTACAAAGACCTTCTCTGTATGTATTAACTTTGCAAGAAAGTCATCCAGGCATGTAC	1429	
Db	402	AspThrSerAsnLeuProAspThrValLeuAsnPheIleArgSerHisProLeuMetAsp	421	
Qy	1430	AATCCAGTGTTCCTATGAACAATCGCCCAATAGTGTATCAAAACGGATGTAATATCA	1489	
Db	422	LysAlaValAsnHisGluHisAsnAsnProValTyrTyrLysArgAspLeu-----Val	439	
Qy	1490	TTTACACAATTTGTCGTAGAC-----CGATGTGATGCAGAAAGTACAGTATGATGTT	1543	
Db	440	PheThrLysLeuValValAspLysIleArgIleAspIleLeuAsnGlnGluTyrIleVal	459	
Qy	1544	ATGTTTATCGGAACAGATGTTGGACCGTCTCTTAAAGTAGTGTTCATTCCTAAGGAGACT	1603	
Db	460	TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln-----	474	
Qy	1604	TGGTATGATTTAGAAGAGTTCTGCTGGAGAAGATGACAGTTTTTCGG-----GAACCG	1657	
Db	475	TyrTyrArgAsnGlyGluSerLeuSerLysLeuLeuAspIlePheGluValAlaProAsn	494	
Qy	1658	ACTGCTATTTTCAGCAATGGAGCTTTCACATAGCAGCAACACTATATATATTTGTTCAACG	1717	
Db	495	GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp	514	
Qy	1718	GCTGGGTGTCCAGCTCCCTTACACCGGTGTGATATTTACGGGAAGCGTGTCTGTAG	1777	
Db	515	HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg	534	
Qy	1778	TGTTGCTCCGCCGAGACCTTACTGTGTGGATGGTGTCTGCA-----TGTTCTCGC	1831	
Db	535	Cys-----ValArgAspProTyrCysGlyTrpAspLysGluAlaAsnThrCysArgPro	552	
Qy	1832	TATTTTCCCACTGCAAGAGACCGCACAGACGACAGATATAGAAATGGAGACCACCTG	1891	
Db	553	Tyr-----GluLeuAspLeuLeuGlnAspValAlaAsn-----	563	
Qy	1892	ACTCACGTGTTACAGATTTACACCATGATAATCACCATGGCCACAGCCCTGAAGAGAGAATC	1951	
Db	564	--GluThrSerAspIleCysAspSerSer-----ValLeuLysLysLysIle	578	
Qy	1952	ATC-----TATGGTGTAGAAATAGTAGCACAATTTTGGAAATGCAGTCCGAAGTCGCAG	2005	
Db	579	ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro	595	
Qy	2006	AGAGCGGTG-----GTCTATTGGCAATTCAGAGGGCGAAATGAAGAGCGAAA	2053	
Db	596	GluValLeuLysAsnGluInValThrTrp---TyrHisHisSerLysAspLysGlyArg	614	
Qy	2054	GAAGAGATCAGATGGGAT-----GATCATATATCATCAGACACATCAAGCGCTTCGTCTA	2107	
Db	615	TyrGluIleArgTyrSerProThrLysTyrIleGluThrThrGluArgGlyLeuValVal	634	
Qy	2108	CGTAGTCTACACAGAGGAGTTCAGGCAATTACTCTGCCAT	2149	

```

Db      635 ValSerValAsnGluAlaAspGlyGlyArgTyrAspCysHis 648
RESULT 13
US-08-835-268-62
; Sequence 62, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-268-62

Alignment Scores:
Pred. No.:          2,07e-69           Length:          724
Score:             785.50              Matches:         221
Percent Similarity: 47.48%             Conservative:    118
Best Local Similarity: 30.95%          Mismatches:     248
Query Match:       16.04%              Indels:        127
DB:                1                   Gaps:          30

US-09-774-490-1 (1-2709) x US-08-835-268-62 (1-724)

Qy      188 AGCGTCGCAGCATGGCTGGTTAACTAGGAATGTCTGTCTTTCTTCCT
Dbs     2 SerLeuLeuGlnLeuSerProLeuLeuAlaLeuLeuLeuLeuLeuCy
Qy      248 ACAGCAAGACCAACTATCAGAAT-----
Dbs     22 GluThrAlaAlaAspTyrGluAsnThrTrpAsnPheTyrTyrGluA
Qy      272 -----GGCAAGAACAATGTGCCAAGGCTGAAATTATCTCTCT
Dbs     42 GlyAsnAspGlnGlyAsnAsnTyrGlyLys-----
Qy      320 GAATCCAACAAATGTGATCACTTTCAATGGCTTGCCCAACAGCTCCAC
```



```
Db 221 ValIlePheArgThrAspLeuTyrAsnThrSerAlaLysArgLeuGluTyrLysPheLys 240
      |||||
QY 833 AGGACAGACGACGATGATTCAGGTGGCTCAATGATCCAAAGTTCATTAGTCCACACCTC 892
      |||||
Db 241 ArgThrLeuLysTyrAspSerLysThrLeuAspLysProAsnPheValGlySerPheAsp 260
      |||||
QY 893 ATCTCAGAGAGTGACAACTCCTGAAGATGACAAAGTATATCTTTCTTCGCGTGAATGCA 952
      |||||
Db 261 IleGlyGlu-----TyrValTyrPhePheArgGluThrAla 273
      |||||
QY 953 ATAGATGGAGAACACTCTCGAAAGCTACTCAGCTAGATAGGTGACAGATATGCAAGAT 1012
      |||||
Db 274 ValGluTyrIleAsnCysGlyLysAlaValTyrSerArgIleAlaArgValCysLysLys 293
      |||||
QY 1013 GACTTTGGAGGCGACAGAGTCTGGTGAATAAATGGACAACTTCTCAAAGCTCCTCTG 1072
      |||||
Db 294 AspValGlyGlyLysAsnLeuLeuAlaHisAsnTTPAlaThrTyrLeuLysAlaArgLeu 313
      |||||
QY 1073 ATTTGCTCAGTCCAGGTCCAAATGGCATTGACACTCATTTTGTAGTAAGTGCAGGATGTA 1132
      |||||
Db 314 AsnCysSerIleSerGlyGlu-----PheProPheTyrPheAsnGluIleGlnSerVal 331
      |||||
QY 1133 TTCCTAATGAACCTTTAAAGATCCTAAATCCAGTTGTATATGGAGTGTTCAGACTTCC 1192
      |||||
Db 332 TyrGlnLeuProSerAspLysSerArg-----PhePheAlaThrPheThrThrSer 348
      |||||
QY 1193 AGTAACATTTTCAAGCGATCAGCGCTGTGTATGTATAGCATGAGTGTGTGAGAAGGGTG 1252
      |||||
Db 349 ThrAsnGlyLeuIleGlySerAlaValCysSerPheHisIleAsnGluIleGlnAlaAla 368
      |||||
QY 1253 TCTCTTGGTCCATATGCCACAGGATGGACCCAACTATCAATGGTGCCT---TATCAA 1309
      |||||
Db 369 PheAsnGlyLysPheLysGluGlnSerSerAsnSerAlaTTPLeuProValLeuAsn 388
      |||||
QY 1310 GGAAGAGTCCCTATCCAGCGCCAGAACTGTCTCCAGCAAAACATTTGGTGGTTTGCAC 1369
      |||||
Db 389 SerArgValProGluProArgProGlyThrCys-----ValAsn 401
      |||||
QY 1370 TCTACAAAGGACCTTCTCTGATGATGTATTAACCTTTGCAAGAGTATCCAGCCCATGTAC 1429
      |||||
Db 402 AspThrSerAsnLeuProAspThrValLeuAsnPheIleArgSerHisProLeuMetAsp 421
      |||||
QY 1430 AATCCAGTGTTCCTATGACAACTCGCCCAATAGTATGATCAAAACCGGATGTAAATATCAA 1489
      |||||
Db 422 LysAlaValAsnHisGluHisAsnAsnProValTyrTyrLysArgAspLeu-----Val 439
      |||||
QY 1490 TTTACACAAATGTCGTAGAC-----CGAGTGGATGCAGAGATGGACAGTATGATGTT 1543
      |||||
Db 440 PheThrLysLeuValValAspLysIleArgIleAspIleLeuAsnGlnGluTyrIleVal 459
      |||||
QY 1544 ATGTTTATCGGAACAGATGTTGGGACCGTCTCTTAAAGTAGTTTCAATTCTTAAGGAGACT 1603
      |||||
Db 460 TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln-----474
      |||||
QY 1604 TGGTATGATTTAGAGAGGTTCTCTGGAAGAAATGACAGTCTTTTCGG-----GAACCG 1657
      |||||
Db 475 TyrTyrArgAsnGlyGluSerLeuSerLysLeuLeuAspIlePheGluValAlaProAsn 494
      |||||
QY 1658 ACTGCTATTTACGACGCTTCTCCTAAGCAGCAACACTATATATTTGGTTCAACG 1717
      |||||
Db 495 GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp 514
      |||||
QY 1718 GCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTACGGGAAAGCGTGTGCTGAG 1777
      |||||
Db 515 HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg 534
      |||||
QY 1778 TGTTCCTCGCCCGCAGACCCCTTACTGTGCTTGGGATGGTCTCTGCA-----TGTTCCTCGC 1831
      |||||
Db 535 Cys-----ValArgAspProTyrCysGlyTTPAspLysGluAlaAsnThrCysArgPro 552
      |||||
QY 1832 TATTTTCCCACTGCAAGAGCGCCACAGACGACAGATATATAAGAAATGGAGCCCACTG 1891
      |||||
```

```
Db 553 Tyr-----GluLeuAspLeuLeuGlnAspValAlaAsn-----563
QY 1892 ACTCACTGTTTCAGACTTACACCATGATTAATCACCATGGCCACAGCCCTGAAGAGAGAATC 1951
      |||||
Db 564 ---GluThrSerAspIleCysAspSer-----ValLeuLysLysLysIle 578
      |||||
QY 1952 ATC-----TATGGGTAGAGAATAGTAGCACATTTTGGAAATGCAGTCCGAAAGTCGCAG 2005
      |||||
Db 579 ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro 595
      |||||
QY 2006 AGAGCGCTG-----GTCTATTGGCAATTCCAGAGCGCAAAATCAAGAGCGCAAAA 2053
      |||||
Db 596 GluValLeuLysAsnGluGlnValThrTrp---TyrHisHisSerLysAspLysGlyArg 614
      |||||
QY 2054 GAAGAGATCAGAGTGGAT-----GATCATATCATCAGACAGATCAAGCCCTTCTGCTA 2107
      |||||
Db 615 TyrGluIleArgTyrSerProThrLysTyrIleGluThrThrGluArgGlyLeuValVal 634
      |||||
QY 2108 CGTAGTCTTACAAACAGAGAGATTTCAGGCAATTACCTCTGCCAT 2149
      |||||
Db 635 ValSerValAsnGluAlaAspGlyArgTyrAspCysHis 648
      |||||
```

Search completed: August 3, 2003, 10:20:43
Job time : 64 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:36 ; Search time 50.5 Seconds
(without alignments)
10317.662 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatctttattttatcatgatg.....agggtttttttcctaataacc 2709

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565704

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09774490/runat_03082003_102927_4471/app_query_fasta_1.2887
-DB=PIR_76 -QFMT=fastan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=BITS -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MOD=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1 1.69 @runat_03082003_102927_4471 -NCPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAEXT=60 -FGAPOP=6
-FGAEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR_76:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	2 D49423	semaphorin III pre
2	82	9.2	666	2 I58169	semaphorin III - m
3	79	8.8	772	2 I48747	semaphorin D - mou
4	58	6.5	772	2 A49069	collapsin - chicke
5	20	2.2	748	2 I48744	semaphorin A - mou
6	15	1.7	749	2 G01856	semaphorin V - hum
7	15	1.7	751	2 I48748	semaphorin E - mou
8	13	1.5	753	2 G02173	semaphorin III fam
9	10	1.1	653	2 T03102	semaphorin homolog
10	10	1.1	834	2 S66498	M-sema F protein p
11	9	1.0	311	2 T11017	MLCL536.34 protein
12	9	1.0	688	1 C1HUS	complement subcomp
13	9	1.0	782	2 I48746	semaphorin C - mou
14	9	1.0	1244	2 S25327	cytoskeleton assem

ryanodine receptor	1.0	4859	2	S74173
ryanodine-binding	1.0	4868	2	B54161
ryanodine receptor	1.0	4869	2	S66572
ryanodine receptor	1.0	4872	2	S27272
ryanodine receptor	1.0	4967	2	S72269
ryanodine receptor	1.0	4969	2	A37113
ryanodine-binding	1.0	5037	1	A54161
ryanodine receptor	1.0	5126	2	S40450
pnd protein - Esch	0.9	50	1	S02763
tryptophan biosynt	0.9	50	2	JU0071
hypothetical prote	0.9	75	2	B38256
hypothetical prote	0.9	101	2	C71047
butyrate-acetoacet	0.9	178	2	AG0244
probable endociti	0.9	223	2	B83932
hypothetical prote	0.9	283	2	D84868
oligopeptide ABC t	0.9	309	2	G87498
oligopeptide ABC t	0.9	326	2	H97346
spore germination	0.9	350	2	C97346
proteinase inhibit	0.9	368	1	I39856
G protein-coupled	0.9	374	2	A59273
type I site-specif	0.9	377	2	B30341
hypothetical prote	0.9	409	2	T09461
hypothetical 53.0K	0.9	421	2	T22969
probable integral	0.9	463	2	J50376
serine-type D-Ala-	0.9	475	2	G95379
B. subtilis negati	0.9	479	2	A64117
B. subtilis negati	0.9	571	2	AB1274
D-lactate dehydrog	0.9	571	2	AC1637
hypothetical prote	0.9	579	2	S51528
phosphate acetyltr	0.9	697	2	T37827
hypothetical prote	0.9	713	2	S50130
semaphorin B - mou	0.9	745	2	T33751
hypothetical prote	0.9	760	2	I48745
probable protein A	0.9	817	2	T51787
semaphorin alpha c	0.9	941	2	A86404
semaphorin F precu	0.9	966	2	A33626
probable ABC trans	0.9	1074	2	JC5928
hypothetical prote	0.9	1146	2	F84487
ryanodine receptor	0.9	2420	2	A84652
ryanodine receptor	0.9	5032	1	A35041
ryanodine receptor	0.9	5035	1	I46646
phycobiliprotein s	0.9	5037	2	B35041
interleukin-7 rece	0.8	10797	2	T30192
hypothetical prote	0.8	16	2	S09700
photosystem I prot	0.8	22	2	B40256
alpha-amylase (EC	0.8	35	2	B82805
hypothetical prote	0.8	37	2	S03674
hypothetical prote	0.8	51	2	S05490
hypothetical prote	0.8	63	2	D42194
hypothetical prote	0.8	68	2	S33581
hypothetical prote	0.8	87	2	A83261
hypothetical prote	0.8	88	2	T25449
hypothetical prote	0.8	88	2	T30626
hemoglobin alpha c	0.8	89	2	I68530
hypothetical prote	0.8	91	2	AF2068
potassium-dependen	0.8	98	2	AE2336
hypothetical prote	0.8	99	2	G84410
hypothetical prote	0.8	99	2	A69514
hypothetical prote	0.8	100	2	AD2566
conserved hypotnet	0.8	102	2	D70408
insulin II precurs	0.8	106	1	IPXL2
hypothetical prote	0.8	106	2	T47814
conserved hypotnet	0.8	106	2	H82421
hypothetical prote	0.8	106	2	B64001
hypothetical prote	0.8	107	2	G72496
hypothetical prote	0.8	110	2	C70599
hypothetical prote	0.8	111	2	T00688
hypothetical 14K p	0.8	118	2	E33958
probable integral	0.8	119	2	F81894
crab protein NMB10	0.8	119	2	C81125
conserved hypotnet	0.8	122	2	F75316
hypothetical prote	0.8	131	2	E75350
hypothetical prote	0.8	131	2	E72469

88 7 0.8 132 2 D72151 B12L protein - var
c 89 7 0.8 135 2 E81433 probable integral
c 90 7 0.8 138 2 D72727 hypothetical prote
c 91 7 0.8 138 2 T18449 pathogenicity fact
c 92 7 0.8 141 2 A32268 carcinomembryonic a
c 93 7 0.8 141 2 I57007 pregnancy-specific
c 94 7 0.8 141 2 C24338 hemoglobin alpha-T
c 95 7 0.8 141 2 T10195 hypothetical prote
c 96 7 0.8 142 2 S71107 cell-cell adhesion
c 97 7 0.8 145 2 T00987 hypothetical prote
c 98 7 0.8 145 2 E90398 hypothetical prote
c 99 7 0.8 147 2 D72504 hypothetical prote
c 100 7 0.8 147 2 A84546 50S ribosomal prot

ALIGNMENTS

RESULT 1

D49423
semaphorin III precursor - human
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: D49423
R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
Cell 75, 1369-1399, 1993
A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
A:Reference number: A49423; MUID:94094332; PMID:8269517
A:Accession: D49423
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-771 <KOL>
A:Cross-references: GB:L26081; NID:g799328; PIDN:AAA65938.1; PID:g436560
C:Genetics:
A:Gene: GDB:SEMA1
A:Cross-references: GDB:283448
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D49423 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTCCTCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGACAATGTGCCAGGCTGAATTCCTCAAGAAGAAATGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGCTGGCCCAAGAGTCCAGTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTGTGGCAAGAGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTATATCAAGATTTTCAAGATTTGTGGCCAGTATCTTACACCAAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAGATGTGCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysIleTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACGCTGTGGACGGGGCTTTTTCATCCAATT 619

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyValaPheHisProIle 140
QY 620 TGCACCTACATTGAATTTGGACATCATCTCAGGACATATATTTTAAGCTGGAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTCTGACGACATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGAGTATTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACTCTTGGCACACACCCATCAGGACAGACAGCATGATTCAGGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCCTCAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLeuSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATCAATAGATGAGAACACATCTCGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACCTAGAAATAGTGCAGATATGCAAGATGACTTTGGAGGCGACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATCTCTCAAAAGCTCGTCTGATTTGCTCAGTGCCAGAGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTCACTCATTTTGTATGAATCACTCAGAGATGTATTCCTTAATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAACATTTTCNAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGAGAAAGGTTGTTCTTGTGTCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACCTTCTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCACTCAGCCATGTACAATCCAGTGTTCCTATGAACAAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTTACAAATTTCTGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAATGACAGTATGATGTTATGTTTATCGGAACAGATGTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAAAGAGTTCCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGGAACCGACTGTCTATTTCAGCAATAGAGCTTTTCACTTAAGCAGACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATTGTTCAACGGCTGGGTTCCTCCAGCTCCCTTTACACCGGTGCTATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTAGCTGTTCCTCCCGAGACCCCTTACTGTGCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTCCCACTCCAAAGAGACGACACAGACGACAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTAGCTTACACCATGATATATCCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAACNTACTATGTTGAGATAGTACACATTTTGGATGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTATTGGCAATTCAGAGCGCAATGAGAGCGAAATGAGAGCAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgGlyGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTAGTGTCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATTACTCTGCTCATCGGTGGAAACATGGTTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGlnHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCTCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGACTTCATGCTAGCTCATCAACACCCCAATCTCAACACATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGAAGAGGACCGAAGCAACAACTCGCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGGAAGCATTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACGAAATTTAGAGGGGACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 2
158169
semaphorin III - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: I58169
R:Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; K
Neuron 14, 949-959, 1995
A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory pr
A:Reference number: 158169; MUID:95267432; PMID:7748562
A:Accession: I58169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-666 <RES>
A:Cross-references: GB:L40484; NID:g703189; PIDN:AAA73934.1; PID:g703190
C:Genetics:

A:Gene: Semaphorin III
C:Superfamily: semaphorin
Alignment Scores:
Pred. No.: 5.44e-79 Length: 666
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.16% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I58169 (1-666)
QY 999 GAGAGTGAATCTCTGAAGATGACAAAGTATATCTTTTCCGTGAAAAATGCAATAGAT 958
Db 128 GluSerAspAsnProGluAspLysValTyrPhePheArgGluAsnAlaIleAsp 147
QY 959 GGAGAACACTCTGAAAAAGCTACTCACGCTAGATAGTTCAGATATGCAAGATGACTTT 1018
Db 148 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 167
QY 1019 GGAGGGCACAGAACTCTGGTGAATAAATGGAACAACATTCTCAAGCTCGTCTGATTTGC 1078
Db 168 GlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCys 187
QY 1079 TCACTGCCAGTCCAAATGCGATTGACACTCATTTTGTGAACTGCAGGATGTATTCCTA 1138
Db 188 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 207
QY 1139 ATGAAC 1144
Db 208 MetAsn 209
RESULT 3
148747
semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48747
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: 148744; MUID:95267431; PMID:7748561
A:Accession: I48747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-772 <RES>
A:Cross-references: EMBL:X85993; NID:g854329; PIDN:CAA59985.1; PID:g854330
C:Genetics:
A:Gene: semD
C:Superfamily: semaphorin
Alignment Scores:
Pred. No.: 9.21e-76 Length: 772
Score: 79.00 Matches: 79
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.83% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I48747 (1-772)
QY 1169 GTATATGAGTGTATTACACTTCCAGTAAACATTTTCAAGGATCAGCGTGTGTATGTAT 1228
Db 324 ValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaValCysMetTyr 343
QY 1229 AGCATGATGTATGTAGAAGGGTGTCTTCTGGTCCATATGCCACAGGATGGACCCAAAC 1288
Db 344 SerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
QY 1289 TATCAATGGTGCCTTATCAAGAGAGTCCCTATCCAGGCCAGCAACTTGTCCCGC 1348
Db 364 TyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383

QY 1349 AAAACATTTGGTGGTTTTCAGCTCTACAAAGGACCTTCTCCTGATGATGCTTATACCTTT 1405
Db 384 LysThrPheGlyClyPheAspSerThrLysAspLeuProAspValIleThrPhe 402

RESULT 4
A49069
semaphorin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
C:Accession: A49069
Cell 75, 217-227, 1993
A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neuron
A:Reference number: A49069; MUID:94006554; PMID:9402908
A:Accession: A49069
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-772 <LUO>
A:Cross-references: GB:U02528; NID:9410078; PIDN:AAC59638.1; PID:g410079
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 4.09e-53 Length: 772
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.48% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A49069 (1-772)

QY 971 GGAAGAGCTACTCAGCTAGATAGTCAGATATCAAGATGACTTTGGAGGCGACAGA 1030
Db 258 GlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArg 277

QY 1031 AGCTGGTGAATAAATGACACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGT 1090
Db 278 SerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGly 297

QY 1091 CCAATGGCATTGACACTCATTTGATGAACCTGACAGGATGATTCCTTAATGAAC 1144
Db 298 ProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 315

RESULT 5
I48744
semaphorin A - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48744
R:Paschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48744
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-748 <RES>
A:Cross-references: EMBL:X85990; NID:9854323; PIDN:CAA59982.1; PID:9854324
C:Genetics:
A:Gene: semA
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 3.94e-12 Length: 748
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48744 (1-748)

QY 1766 CGGTGCTGCTAGTGTTCCTCCGACGACCTTACTGTGCTGGATGGTTCATCT 1825
Db 521 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySerAlaCys 540

RESULT 6
G01856
semaphorin V - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G01856
R:Sekido, Y.
submitted to the EMBL Data Library, June 1995
A:Reference number: G08634
A:Accession: G01856
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-749 <SEK>
A:Cross-references: EMBL:U28369; NID:974283; PIDN:AAD09138.1; PID:g974284
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 9.73e-07 Length: 749
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x G01856 (1-749)

QY 1028 AGAAGTCTGGTGAATAAATGACACATTCCTCAAGCTCGTCTG 1072
Db 277 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 291

RESULT 7
I48748
semaphorin E - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48748
R:Paschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48748
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-751 <RES>
A:Cross-references: EMBL:X85994; NID:9854331; PIDN:CAA59986.1; PID:9854332
C:Genetics:
A:Gene: semE
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 9.73e-07 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48748 (1-751)

QY 1028 AGAAGTCTGGTGAATAAATGACACATTCCTCAAGCTCGTCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 8
G02173
semaphorin III family homolog - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G02173

R.Naylor, S.
submitted to the EMBL Data Library, October 1995
A:Reference number: G09275
A:Accession: G02173
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-753 <NAY>
A:Cross-references: EMBL:U38276; NID:g1061350; PIDN:AAB18276.1; PID:g1061351
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 0.00014 Length: 753
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x G02173 (1-753)

QY 1778 TGTTCCTCGCCGAGACCTTACTGTGCTTGGGATGGT 1816
|||||
Db 526 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 538
|||||

RESULT 9
T03102
semaphorin homolog A3 - alcelaphine herpesvirus 1
C:Species: alcelaphine herpesvirus 1
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
A:Accession: T03102
R:Jensner, A.; Pflanz, R.; Fleckenstein, B.
J. Virol. 71, 6517-6525, 1997
A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.
A:Reference number: Z14840; MUID:97404659; PMID:9261371
A:Accession: T03102
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-653 <ENS>
A:Cross-references: EMBL:AF005370; NID:g2337967; PIDN:AAC58054.1; PID:g2337970

Alignment Scores:
Pred. No.: 0.244 Length: 653
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T03102 (1-653)

QY 1043 AAATGGACACATTCCTCAAGCTCGTCTG 1072
|||||
Db 302 LysTrpThrPheLeuLysAlaArgLeu 311
|||||

RESULT 10
S66498
M-sema F protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999
A:Accession: S66498
R:Inagaki, S.; Furuyama, T.; Iwahaishi, Y.
FEBS Lett. 370, 269-272, 1995
A:Title: Identification of a member of mouse semaphorin family.
A:Reference number: S66498; MUID:95385809; PMID:7656991
A:Accession: S66498
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-834 <INA>
A:Cross-references: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599
C:Superfamily: semaphorin
F1-21/Domain: signal sequence #status predicted <STO>
F122-834/Product: M-sema F protein #status predicted <MAT>

Alignment Scores:
Pred. No.: 0.237 Length: 834
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S66498 (1-834)

QY 1043 AAATGGACACATTCCTCAAGCTCGTCTG 1072
|||||
Db 274 LysTrpThrPheLeuLysAlaArgLeu 283
|||||

RESULT 11
T11017

MLCL536.34 protein - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 23-Mar-2001
A:Accession: T11017; S72779
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z16918
A:Accession: T11017
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-311 <PAR>
A:Cross-references: EMBL:Z99125; NID:g2398683; PIDN:CAB16178.1; PID:g2398713
R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B1496.
A:Reference number: S72695
A:Accession: S72779
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'M', 90-311 <SMI>
A:Cross-references: EMBL:U00013; NID:g466868; PIDN:AAAL7144.1; PID:g466899
C:Genetics:
A:Gene: MLCL536.34
A:Start codon: GTG

Alignment Scores:
Pred. No.: 3.17 Length: 311
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T11017 (1-311)

QY 516 GCCCACTTGCATTCATCTCTCTGTG 490
|||||
Db 225 AlaHisLeuHisSerSerLeuLeuVal 233
|||||

RESULT 12
CIHUS

complement subcomponent C 1S8AR.GIF (EC 3.4.21.42) precursor [validated] - human
N:Alternate names: C1 esterase precursor
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 03-Jun-2002
A:Accession: A40496; A27381; S00224; S26732; S05634; A05140; A25396; A38407; B37820
R:Kusumoto, H.; Hirose, S.; Sallier, J.P.; Hagen, F.S.; Kurachi, K.
Proc. Natl. Acad. Sci. U.S.A. 85, 7307-7311, 1988
A:Title: Human genes for complement components C1r and C1s in a close tail-to-tail arrangement
A:Reference number: A40496; MUID:89017187; PMID:2459702
A:Accession: A40496
A:Molecule type: mRNA
A:Residues: 1-688 <KUS>
A:Cross-references: GB:J04080; NID:g179645; PIDN:AAAS1852.1; PID:g179646
R:Tosi, M.; Duponchel, C.; Meo, T.; Jullier, C.
Biochemistry 26, 8516-8524, 1987
A:Title: Complete cDNA sequence of human complement C1s and close physical linkage of th

```

F;16-437/Product: complement subcomponent C1s chain A (heavy chain) #status experimental
F;135-171/Domain: EGF homology <EGF>
F;175-287/Domain: C1r/C1s repeat homology <C1R2>
F;259-354/Domain: complement factor H repeat homology <FH1>
F;359-421/Domain: complement factor H repeat homology <FH2>
F;438-688/Product: complement subcomponent C1s chain B (light chain) #status experimental
F;438-675/Domain: trypsin homology <TRY>
F;65-83,135-147,143-156,158-171,175-202,234-251,294-341,321-354,359-403,386-421,425-549
F;149/Modified site: erythro-beta-hydroxyasparagine (Asn) (partial) #status experimental
F;174,406/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;437-438/Cleavage site: Arg-Ile (complement subcomponent C1r) #status experimental
F;475,529,632/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.: 2,9 Length: 688
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x CIHUS (1-688)

Qy 151 TTGTGCGCCAGAGAGTTCAACAACAT 125
Db 82 LeuCyseGlyGlnArgSerSerAsnAsn 90

RESULT 13
I48746
semaphorin C - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 03-Nov-2000
C;Accession: I48746
R;Buschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A;Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A;Reference number: I48744; MUID:95267431; PMID:7748561
A;Accession: I48746
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-782 <RES>
A;Cross-references: EMBL:X85992; NID:g854327; PIDN:CAAS9984.1; PID:g854328
C;Genetics:
A;Gene: semC
C;Superfamily: semaphorin

Alignment Scores:
Pred. No.: 2,86 Length: 782
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48746 (1-782)

Qy 1784 CTCGCCCGACACCTTACTGCTTGG 1810
Db 483 LeuAlaAArgAspProtyrCysAlaTrp 491

RESULT 14
S25327
cytoskeleton assembly control protein SLAI - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YBL007c; protein YBL0321
C;Species: Saccharomyces cerevisiae
C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 29-Oct-1999
C;Accession: S25327; A40673; S45740; S36355; S42277
R;Delaveau, T.; Jacq, C.; Perea, J.
Yeast 8, 761-768, 1992
A;Title: Sequence of a 12.7 kb segment of yeast chromosome II identifies a PDR-like gene
A;Reference number: S25326; MUID:93070613; PMID:1441753
A;Accession: S25327
A;Molecule type: DNA

```


F:4471-4494/Domain: transmembrane #status predicted <TM6>
F:4611-4642/Domain: transmembrane #status predicted <TM7>
F:4659-4678/Domain: transmembrane #status predicted <TM8>
F:4701-4720/Domain: transmembrane #status predicted <TM9>
F:4736-4759/Domain: transmembrane #status predicted <TM10>
F:130,290,1243/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent kinases)
F:2706/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinases)
F:4686/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,34 Length: 4859
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S74173 (1-4859)

Qy 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2132 ArgGlySerThrProLeuAspValAla 2140

RESULT 16
B54161
ryanodine-binding protein beta form - bullfrog
C/Species: Rana catesbeiana (bullfrog)
C/Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 20-Jun-2000
C/Accession: B54161
R/Oyamada, H.; Mucayama, T.; Takagi, T.; Iino, M.; Iwabe, N.; Miyata, T.; Ogawa, Y.; Endo, J. Biol. Chem. 269, 17206-17214, 1994
A/Title: Primary structure and distribution of ryanodine-binding protein isoforms of the A/Reference number: A54161; MUID:94274714; PMID:8006029
A/Accession: B54161
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-4868 <OYA>
A/Cross-references: GB:D21071; NID:g1856973; PIDN:BAA04647.1; PID:G538247
C/Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2,34 Length: 4868
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B54161 (1-4868)

Qy 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 17
S66572
ryanodine receptor type 3 - chicken
C/Species: Gallus gallus (chicken)
C/Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C/Accession: S66572
R/Ottini, L.; Marziali, G.; Conti, A.; Charlesworth, A.; Sorrentino, V. Biochem. J. 315, 207-216, 1996
A/Title: Alpha and beta isoforms of ryanodine receptor from chicken skeletal muscle are A/Reference number: S66572; MUID:96207583; PMID:8670108
A/Accession: S66572
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-4869 <OTT>
A/Cross-references: EMBL:X95267; NID:g1212911; PIDN:CAA64563.1; PID:g1212912
C/Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2,34 Length: 4869

Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S66572 (1-4869)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
Db 2135 ArgGlySerThrProLeuAspValAla 2143

RESULT 18
S27272
ryanodine receptor, brain - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jun-1999
C:Accession: S27272
R:Hakamata, Y.; Nakai, J.; Takeshima, H.; Imoto, K.
FBS Lett. 312, 229-235, 1992
A:Title: Primary structure and distribution of a novel ryanodine receptor/calcium release
A:Reference number: S27272; MUID:93050200; PMID:1330694
A:Accession: S27272
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4872 <HAK>
A:Cross-references: EMBL:X68650
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo

Alignment Scores:
Pred. No.: 2.34 Length: 4872
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S27272 (1-4872)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
Db 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 19
S27269
ryanodine receptor isoform 2, cardiac muscle - human
N:Alternate names: calcium release channel protein isoform 2
C:Species: Homo sapiens (man)
C>Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 20-Jun-2000
C:Accession: S72269
R:Tunwell, R.E.A.; Wickenden, C.; Bertrand, B.M.A.; Shevchenko, V.I.; Walsh, M.B.; Allen
Biochem. J. 318, 477-487, 1996
A:Title: The human cardiac muscle ryanodine receptor-calcium release channel: identifica
A:Reference number: S72269; MUID:96404895; PMID:8809036
A:Accession: S72269
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-4967 <TUN>
A:Cross-references: EMBL:X98330; NID:g1526977; PIDN:CA66975.1; PID:g1526978
A:Experimental source: cardiac muscle; sarcoplasmic reticulum
C:Genetics:
A:Gene: RyR-2
C:Complex: homotetramer
C:Function:
A:Description: mediates rapid release of Ca2+ ions from intracellular stores during card
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
C:Keywords: alternative splicing; calcium channel; cardiac muscle; glycoprotein; heart;
F:710-715/Region: adenine nucleotide binding #status predicted
F:1148-1153/Region: adenine nucleotide binding #status predicted
F:2338-2342/Region: adenine nucleotide binding #status predicted
F:2626-2631/Region: adenine nucleotide binding #status predicted
F:4030-4048/Domain: calcium binding #status predicted <CAL>
F:4499-4519/Domain: transmembrane #status predicted <TM01>

F:4572-4593/Domain: transmembrane #status predicted <TM02>
F:4719-4744/Domain: transmembrane #status predicted <TM03>
F:4767-4789/Domain: transmembrane #status predicted <TM04>
F:4809-4829/Domain: transmembrane #status predicted <TM05>
F:4844-4867/Domain: transmembrane #status predicted <TM06>
F:140-301/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent kin
F:2808/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status
F:4794/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2.34 Length: 4967
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S72269 (1-4967)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
Db 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 20
A37113
ryanodine receptor, cardiac muscle - rabbit
N:Alternate names: ryanodine receptor 2
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 20-Jun-2000
C:Accession: A37113; S74212
J:Otsu, K.; Willard, H.F.; Khanna, V.K.; Zorzato, F.; Green, N.M.; MacLennan, D.H.
J. Biol. Chem. 265, 13472-13483, 1990
A:Title: Molecular cloning of cDNA encoding the Ca(2+) release channel (ryanodine recept
A:Reference number: A37113; MUID:90337947; PMID:2380170
A:Accession: A37113
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4969 <OTS>
A:Cross-references: GB:M59743; GB:J05564; NID:g164831; PIDN:AAA31179.1; PID:g164832
R:Nishida, K.; Otsu, K.; Hori, M.; Kuzuya, T.; Tada, M.
Eur. J. Biochem. 240, 408-415, 1996
A:Title: Cloning and characterization of the 5'-upstream regulatory region of the Ca2+-r
A:Reference number: S74212; MUID:96439071; PMID:8841406
A:Accession: S74212
A:Molecule type: DNA
A:Residues: 1-16 <NIS>
A:Cross-references: EMBL:X99486
A:Experimental source: strain New Zealand White
C:Genetics:
A:Gene: RYR2
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
C:Keywords: cardiac muscle; heart; phosphoprotein; transmembrane protein

Alignment Scores:
Pred. No.: 2.34 Length: 4969
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A37113 (1-4969)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
Db 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 21
A54161
ryanodine-binding protein alpha form - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A54161

R.; Oyamada, H.; Murayama, T.; Takagi, T.; Iino, M.; Iwabe, N.; Miyata, T.; Ogawa, Y.; Endo, J. *Biol. Chem.* 269, 17206-17214, 1994
A:Title: Primary structure and distribution of ryanodine-binding protein isoforms of the A:Reference number: A54161; MUID:94274714; PMID:8006029
A:Accession: A54161
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-5037 <OVA>
A:Cross-references: GB:D21070; NID:g1856972; PIDN:BAA04646.1; PID:g538246
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2.33 Length: 5037
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x A54161 (1-5037)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2253 ArgGlySerThrProLeuAspValAla 2261
|||||

RESULT 22
S40450
ryanodine receptor/calcium release channel - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 11-Jun-1999
C:Accession: S40450
R:Takekuma, H.; Nishi, M.; Iwabe, N.; Miyata, T.; Hosoya, T.; Masai, I.; Hotta, Y. *FEBS Lett.* 337, 81-87, 1994
A:Title: Isolation and characterization of a gene for a ryanodine receptor/calcium release channel
A:Reference number: S40450; MUID:94102409; PMID:8276118
A:Accession: S40450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5126 <TAK>
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2.33 Length: 5126
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S40450 (1-5126)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2355 ArgGlySerThrProLeuAspValAla 2363
|||||

RESULT 23
S02763
pnd protein - *Escherichia coli* plasmid R16
C:Species: *Escherichia coli*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S02763
R:Sakikawa, T.; Akimoto, S.; Ohnishi, Y. *Biochim. Biophys. Acta* 1007, 158-166, 1989
A:Title: The pnd gene in *E. coli* plasmid R16: nucleotide sequence and gene expression level
A:Reference number: S02763; MUID:89150247; PMID:2465777
A:Accession: S02763
A:Molecule type: DNA
A:Residues: 1-50 <SAK>
A:Cross-references: EMBL:X12833; NID:942453; PIDN:CAA31320.1; PID:g42454
C:Genetics:
A:Gene: pnd
A:Genome: plasmid
C:Superfamily: flmA protein

Alignment Scores:
Pred. No.: 46.4 Length: 50
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x S02763 (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
|||||
Db 3 GlnArgThrPheLeuMetMetLeu 10
|||||

RESULT 24
JU0071
pnd protein - *Escherichia coli* plasmid R483
C:Species: *Escherichia coli*
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Jun-2000
C:Accession: JU0071
R:Ono, K.; Akimoto, S.; Ohnishi, Y. *Microbiol. Immunol.* 31, 1071-1083, 1987
A:Title: Nucleotide sequence of the pnd gene in plasmid R483 and role of the pnd gene product in the degradation of stable RNA in growing *Escherichia coli*
A:Reference number: JU0071; MUID:88174457; PMID:3328043
A:Accession: JU0071
A:Molecule type: DNA
A:Residues: 1-50 <ONO>
A:Cross-references: GB:D00364; NID:g216610; PIDN:BAA00270.1; PID:g216612
C:Comment: The protein increases the degradation of stable RNA in growing *Escherichia coli*
C:Genetics:
A:Genome: plasmid
C:Superfamily: flmA protein

Alignment Scores:
Pred. No.: 46.4 Length: 50
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x JU0071 (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
|||||
Db 3 GlnArgThrPheLeuMetMetLeu 10
|||||

RESULT 25
B38256
tryptophan biosynthetic enzyme transcription regulatory protein mtrB [validated] - *Bacillus subtilis*
N:Alternate names: trp RNA-binding attenuation protein; tryptophan operon RNA-binding protein
C:Species: *Bacillus subtilis*
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 15-Sep-2000
C:Accession: B38256; H69661
R:Gollnick, P.; Ishino, S.; Kuroda, M.I.; Henner, D.J.; Yanofsky, C. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8726-8730, 1990
A:Title: The mtr locus is a two-gene operon required for transcription attenuation in *E. coli*
A:Reference number: A38256; MUID:91062353; PMID:2123343
A:Accession: B38256
A:Molecule type: DNA
A:Residues: 1-75 <COL>
A:Cross-references: GB:M37320; NID:g143230; PIDN:AAA22616.1; PID:g143234
R:Antson, A.A.; Dodson, E.J.; Gollnick, P. *Submitted to the Brookhaven Protein Data Bank, February 1995*
A:Reference number: A66916; PDB:1WAP
A:Contents: annotation; X-ray crystallography, 1.80 angstroms
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrarini, E. *Nature* 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. *tech. J.*

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Saco, I.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tanakara, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: H69661

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-75 <XUN>

A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14193.1; PID:92634495

A:Experimental source: strain 168

C:Comment: This protein is required for transcription termination in the trp leader region

C:Genetics:

A:Gene: mtrB

C:Keywords: RNA binding; transcription termination

Alignment Scores:
Pred. No.: 44.4 Length: 75
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B38256 (1-75)

QY 2034 GCGGAATGAAGAGCGCAAGAG 2057

Db 68 GlyGluMetLysSerGluLysLys 75

RESULT 26

C71047

C:Species: *Pyrococcus horikoshii*

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C:Accession: C71047

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: C71047

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-101 <KAW>

A:Cross-references: GB:AP000006; NID:93236133; PIDN:BAA30779.1; PID:93258096

A:Experimental source: strain OT3

A>Note: This accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH1667

C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1667

Alignment Scores:
Pred. No.: 43 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x C71047 (1-101)

QY 67 CGTGGTGCACACCCCTCGACGTG 44

Db 24 ArgGlySerThrProLeuAspVal 31

RESULT 27

AG0244

hypothetical protein YPO2006 [imported] - *Yersinia pestis* (strain CO92)

C:Species: *Yersinia pestis*

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AG0244

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell.

Nature 413, 523-527, 2001

A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0244

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-178 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90819.1; PID:915980020; GSPDB:GN00175

C:Genetics:

A:Gene: YPO2006

Alignment Scores:
Pred. No.: 40.4 Length: 178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AG0244 (1-178)

QY 2517 GCTCAGACATCTCGGTGCCCTC 2494

Db 138 AlaGlnThrLeuLeuGlyAlaLeu 145

RESULT 28

B83932

butyrate-acetoacetate CoA-transferase B82258 [imported] - *Bacillus halodurans* (strain C-

C:Species: *Bacillus halodurans*

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: B83932

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B83932

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-223 <STO>

A:Cross-references: GB:AP001514; GB:BA000004; NID:910174613; PIDN:BA805977.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: B82258

C:Superfamily: 3-oxoadipate CoA-transferase beta chain; 3-oxoadipate CoA-transferase bet

Alignment Scores:
Pred. No.: 39.4 Length: 223
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B83932 (1-223)

QY 2120 GTTGTAGACTACGTAGCAGAGGC 2097

Db 36 ValValAspTyrValAlaGluGly 43

RESULT 29

D84868

probable endochitinase [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: D84868

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84868

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-283 <STO>

A:Cross-references: GB:AE002093; NID:G2281108; PIDN:AAB64044.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g43620

A:Map position: 2

C:Superfamily: lectin-related plant chitinase; hevvin chitin-binding domain homology; pl

Alignment Scores:

Pred. No.:	38.4	Length:	283
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x D84868 (1-283)

QY 2319 TCACGCTCATCAACCAACCCCAATC 2342

DB 64 CysSerSerSerThrThrProile 71

RESULT 30

GB7498

hypothetical protein CC2012 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: GB7498

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: GB7498

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-309 <STO>

A:Cross-references: GB:AE005673; NID:gl3423483; PIDN:AAK23987.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2012

Alignment Scores:

Pred. No.:	38	Length:	309
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.92%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x GB7498 (1-309)

QY 1074 ATCAGACGAGCTTTCAGGAATGTT 1051

DB 38 fIeArgArgAlaLeuArgAsnVal 45

RESULT 31

H97346

oligopeptide ABC transporter, ATPase component CAC3641 [imported] - Clostridium acetobut C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: H97346

R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: H97346

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81563.1; PID:gl5026741; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3641

Alignment Scores:

Pred. No.:	37.8	Length:	326
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x H97346 (1-326)

QY 296 AAATTATCTACAAAGAAATGTTG 319

DB 84 LysLeuSerTyrLysGluMetLeu 91

RESULT 32

C97346

oligopeptide ABC transporter, ATPase component CAC3635 [imported] - Clostridium acetobut C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 17-May-2002

C:Accession: C97346

R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97346

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81558.1; PID:gl5026736; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3635

C:Superfamily: inner membrane protein malk; ATP-binding cassette homology

Alignment Scores:

Pred. No.:	37.5	Length:	350
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x C97346 (1-350)

QY 296 AAATTATCTACAAAGAAATGTTG 319

DB 108 LysLeuSerTyrLysGluMetLeu 115

RESULT 33

I39856

spore germination apparatus protein gerBB - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: I39856; A69630

R:Corfe, B.M.; Sammons, R.L.; Smith, D.A.; Mauel, C. Microbiology 140, 471-478, 1994

A:Title: The gerB region of the Bacillus subtilis 168 chromosome encodes a homologue of

A:Reference number: I39854; MUID:94282292; PMID:8012571

A:Accession: I39856

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-368 <RES>

Alignment Scores:

Pred. No.:	37.5	Length:	350
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x C97346 (1-350)

QY 296 AAATTATCTACAAAGAAATGTTG 319

DB 108 LysLeuSerTyrLysGluMetLeu 115

RESULT 33

I39856

spore germination apparatus protein gerBB - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: I39856; A69630

R:Corfe, B.M.; Sammons, R.L.; Smith, D.A.; Mauel, C. Microbiology 140, 471-478, 1994

A:Title: The gerB region of the Bacillus subtilis 168 chromosome encodes a homologue of

A:Reference number: I39854; MUID:94282292; PMID:8012571

A:Accession: I39856

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-368 <RES>

A:Cross-references: GB:L16960; NID:G289274; PIDN:AAA22467.1; PID:G289276
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akouchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, H.F.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377

A:Accession: A69630
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-368 <KUN>
 A:Cross-references: GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15598.1; PID:G2636107
 A:Experimental source: strain 168
 A:Note: Germination response to the combination of glucose, fructose, L-asparagine, and
 C:Genetics:
 A:Gene: gerBB; gerB-2
 A:Map position: 314 deg.
 C:Superfamily: spore germination protein

Alignment Scores:
 Pred. No.: 37.3 Length: 368
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x I39856 (1-368)

QY 2481 GTCCTCCTGTTCTACCTTCCTTA 2458
 Db 56 ValLeuLeuPheLeuProPheLeu 63
 RESULT 34
 A59273
 Proteinase inhibitor 8 - human
 N:Alternate names: cytoplasmic antiproteinase 2 (CAP-2)
 C:Species: Homo sapiens (man)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: A59273
 R:Spracher, C.A.; Morgenstern, K.A.; Mathewes, S.; Dahlen, J.R.; Schrader, S.K.; Foster,
 J. Biol. Chem. 270, 29854-29861, 1995
 A:Title: Molecular cloning, expression, and partial characterization of two novel member
 A:Reference number: A59273; MUID:96102039; PMID:8530382
 A:Accession: A59273
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-374 <SPR>

A:Cross-references: GB:L40377; NID:gl160926; PIDN:AC41939.1; PID:gl160927
 A:Experimental source: tissue type placenta; note (vector lambda gt11); gene CAP2
 C:Genetics:
 A:Gene: GDB:P18; CAP2
 A:Cross-references: GDB:599392; OMIM:601697
 A:Map position: 18q21.3-18q21.3
 C:Superfamily: antithrombin III
 C:Keywords: proteinase inhibitor
 F:339/Inhibitory site: Arg (unidentified proteinase) #status predicted

Alignment Scores:
 Pred. No.: 37.2 Length: 374
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0
 US-09-774-490-1 (1-2709) x A59273 (1-374)

QY 1590 TTCCTAAGGAGACTTGGTATGATT 1613

Db 282 PheLeuArgArgLeuGlyMetIle 289

RESULT 35

B30341

G protein-coupled receptor RDC4 - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999

C:Accession: B30341; S12822

R:Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, C.; Simo

Science 244, 569-572, 1999

A:Title: Selective amplification and cloning of four new members of the G protein-coupled

A:Reference number: A30341; MUID:89242119; PMID:2541503

A:Accession: B30341

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-377 <LIB>

A:Cross-references: EMBL:X14049; NID:G900; PIDN:CAA32207.1; PID:G901

R:Libert, F.; Parmentier, M.; Lefort, A.; Dumont, J.E.; Vassart, G.

Nucleic Acids Res. 18, 1916, 1990

A:Title: Complete nucleotide sequence of a putative G protein coupled receptor: RDC4.

A:Reference number: S12822; MUID:90245610; PMID:2159630

A:Accession: S12822

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-377 <L12>

A:Cross-references: EMBL:X14049; NID:G900; PIDN:CAA32207.1; PID:G901

C:Genetics:

A:Gene: RDC4

C:Superfamily: octopamine receptor type I

C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Alignment Scores:
 Pred. No.: 37.2 Length: 377
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B30341 (1-377)

QY 1917 ATATACCATGGCCACAGCCCTG 1940

Db 47 IleIleThrMetAlaThrAlaLeu 54

RESULT 36

T09461

type I site-specific deoxyribonuclease (EC 3.1.21.3) LldI chain HsdS - Lactococcus lacti

C:Species: Lactococcus lactis subsp. lactis

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T09461

R:Deng, Y.M.; Dunn, N.W.

A:Description: LldI, a type I restriction modification system in lactococcus lactis biov

A:Reference number: 216679

A:Accession: T09461

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-409 <DEN>

A:Cross-references: EMBL:AF034786; NID:G2689697; PID:G2689700

A:Experimental source: biovar. diacetylactis; strain UK19161

C:Genetics:

A:Gene: hsdS

A:Genome: plasmid pND861

C:Keywords: hydrolase

Alignment Scores:
Pred. No.: 36.8 Length: 409
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T09461 (1-409)

QY 1135 GAATACATCTCTGAGTTTCATCAA 1112

Db 117 GluTyrIleLeuGlnPheIleIlys 124

RESULT 37

T22969

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22969

R:Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19644

A:Accession: T22969

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-421 <WIL>

A:Cross-references: EMBL:Z81557; PIDN:CAB04538.1; GSPDB:GN000023; CESP:F59A1.13

A:Experimental source: clone F59A1

C:Genetics:

A:Gene: CESP:F59A1.13

A:Map position: 5

A:Introns: 27/1; 116/1; 245/3; 286/3; 340/3; 381/3

Alignment Scores:

Pred. No.: 36.7 Length: 421
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T22969 (1-421)

QY 291 CTGGCACATGTTCTTCCCATTC 268

Db 77 LeuGlyThrLeuPhePheProPhe 84

RESULT 38

JS0376

C:Species: soybean chlorotic mottle virus

A:Note: host Glycine max (soybean)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Jul-2000

C:Accession: JS0376

R:Hasegawa, A.; Verver, J.; Shimada, A.; Saito, M.; Goldbach, R.; van Kammen, A.; Miki,

Nucleic Acids Res. 17, 993-1003, 1989

A:Title: The complete sequence of soybean chlorotic mottle virus DNA and the identification

A:Reference number: JS0372; MUID:90098857; PMID:2602148

A:Accession: JS0376

A:Molecule type: DNA

A:Residues: 1-463 <HAS>

A:Cross-references: GB:X15828; NID:G58833; PIDN:CAA33831.1; PID:G58840

C:Superfamily: soybean chlorotic mottle virus hypothetical 53.0K protein

Alignment Scores:

Pred. No.: 36.3 Length: 463
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x JS0376 (1-463)

QY 1338 CTTGTCCAGCAAAACATTTGGTG 1361

Db 303 LeuValProAlaLysHisLeuVal 310

RESULT 39

G95379

probable integral membrane transporter SMa1717 [imported] - Sinorhizobium meliloti (stra

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: G95379

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, K.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: G95379

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-475 <KUR>

A:Cross-references: GB:AB006469; PIDN:AAK65601.1; PID:G14524083; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, I.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMa1717

A:Genome: plasmid

C:Superfamily: conserved hypothetical protein HI0125

Alignment Scores:
Pred. No.: 36.2 Length: 475
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x G95379 (1-475)

QY 2478 CTCCTGTTTCTACCTTTCTTTATTT 2455

Db 370 LeuLeuPheLeuProPheLeuPhe 377

RESULT 40

A64117

serine-type D-Ala-D-Ala carboxypeptidase homolog - Haemophilus influenzae (strain Rd KX2

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-2000

C:Accession: A64117

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.B.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: A64117

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-479 <TIGR>

A:Cross-references: GB:U32812; GB:U32812; NID:G1574784; PIDN:AAC22975.1; PID:G1574789; T

C:Superfamily: D-alanyl-D-alanine carboxypeptidase

Alignment Scores:

Pred. No.: 36.2 Length: 479

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A64117 (1-479)

QY 1951 GATTCTCTCTTCAGGGCTGCGCC 1928

DB 316 AspSerLeupheArgAlaValAla 323

RESULT 41

AB1274

B. subtilis negative regulator of PtsZ ring formation (EzrA) homolog lmo1594 [imported]

C:Species: *Listeria monocytogenes*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1274

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Krefit, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1274

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-571 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99672.1; PID:g16411023; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1594

Alignment Scores:

Pred. No.: 35.5 Length: 571
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AB1274 (1-571)

QY 1610 GATTCTCTCTTCAGGGCTGCGAA 1633

DB 82 AspLeuGluValLeuLeuGlu 89

RESULT 42

AB1637

B. subtilis negative regulator of PtsZ ring formation (EzrA) homolog lin1636 [imported]

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AC1637

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Krefit, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1637

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-571 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96867.1; PID:g16414123; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin1636

Alignment Scores:

Pred. No.: 35.5 Length: 571
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AC1637 (1-571)

QY 1610 GATTCTCTCTTCAGGGCTGCGAA 1633

DB 82 AspLeuGluValLeuLeuGlu 89

RESULT 43

SS1528

D-lactate dehydrogenase (cytochrome) (EC 1.1.2.4) - yeast (*Kluyveromyces marxianus* var.

C:Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphaerica*

C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999

C:Accession: SS1528

R:Lodi, T.; O'Connor, D.; Goffrini, P.; Ferrero, I.

Mol. Gen. Genet. 244, 622-629, 1994

A:Title: Carbon catabolite repression in *Kluyveromyces lactis*: isolation and characteriz

A:Reference number: SS1528; MUID:95059916; PMID:7969031

A:Accession: SS1528

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-579 <LOD>

A:Cross-references: EMBL:X71628; NID:g602028; PIDN:CAA50635.1; PID:g602029

A:Note: the source is designated as *Kluyveromyces lactis*

C:Keywords: oxidoreductase

Alignment Scores:

Pred. No.: 35.5 Length: 579
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x SS1528 (1-579)

QY 495 GAAGAGATGAATCACTGGCTG 518

DB 470 GluGluMetAenAlaSerGlyLeu 477

RESULT 44

T37827

hypothetical protein SPAC17A5.12 - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T37827

R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21737

A:Accession: T37827

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-697 <DEV>

A:Cross-references: EMBL:Z98849; PIDN:CAB11512.1; GSPDB:GN00066; SPDB:SPAC17A5.12

A:Experimental source: strain 972h-; cosmid c17A5

C:Genetics:

A:Gene: SPDB:SPAC17A5.12

A:Map position: 1

Alignment Scores:

Pred. No.: 34.7 Length: 697
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T37827 (1-697)

QY 383 TTGGATGAGGAACGGAGTAGGCTG 406
DB |||||||||||||||||||||

QY 592 LeuAspGluGluArgSerArgLeu 599
DB |||||||||||||||||||||

RESULT 45

S50130
phosphate acetyltransferase (EC 2.3.1.8) [validated] - Escherichia coli (strain K-12)
N:Alternate names: phosphotransacetylase
C:Species: Escherichia coli
C:Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C:Accession: S50130
R:Masuyama, A.; Yamamoto-Otake, H.; Hewitt, J.; MacGillivray, R.T.A.; Nakano, E.
Biochim. Biophys. Acta 1219, 559-562, 1994
A:Title: Nucleotide sequence of the phosphotransacetylase gene of Escherichia coli strain K-12
A:Reference number: S50130; MUID:95002178; PMID:7918659
A:Accession: S50130
A:Molecule type: DNA
A:Residues: 1-713 <MAT>
A:Cross-references: EMBL:D21123; NID:9577281; PIDN:BAA04663.1; PID:G601935
A:Experimental source: strain K-12 1100
A:Note: the authors translated the codon CGT for residue 113 as Glu, GTA for residue 407
C:Genetics: A:Note: part of this sequence, including the amino end, was confirmed by protein sequencing
A:Gene: pta
A:Start codon: GTG
C:Function:
A:Description: EC 2.3.1.8 [validated, MUID:95002178]
C:Superfamily: phosphate acetyltransferase pta
C:Keywords: acyltransferase; coenzyme A
F.2-713/Product: phosphate acetyltransferase #status experimental <MAT>

Alignment Scores:
Pred. No.: 34.7 Length: 713
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S50130 (1-713)

QY 1788 GCGAGCGACACTGACGACGCT 1765
DB |||||||||||||||||||||

QY 260 AlaArgGlnHisSerAlaHisAla 267
DB |||||||||||||||||||||

RESULT 46

T33751
hypothetical protein R11E3.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T33751
R:Langston, Z.; Wohldmann, P.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid R11E3.
A:Reference number: Z21397
A:Accession: T33751
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-745 <LAN>
A:Cross-references: EMBL:AF100669; PIDN:AC68993.1; GSPDB:GN00022; CESP:R11E3.8
A:Experimental source: strain Bristol N2; clone R11E3
C:Genetics:
A:Map position: 4
A:Gene: CESP:R11E3.8
A:Introns: 104/2; 209/3; 379/3; 528/2; 576/3; 711/2
C:Superfamily: acylaminoacyl-peptidase

Alignment Scores:
Pred. No.: 34.5 Length: 745
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T33751 (1-745)

QY 1362 CCACCAATGTTTTCGCGGACAA 1339
DB |||||||||||||||||||||

QY 447 PropRoanValLeuLeuGlyGln 454
DB |||||||||||||||||||||

RESULT 47

I48745
semaphorin B - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C:Accession: I48745
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates a guidance cue for retinal ganglion cell axons
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48745
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-760 <RES>
A:Cross-references: EMBL:X85991; NID:9854325; PIDN:CAA59983.1; PID:9854326
C:Genetics:
A:Gene: semB
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 34.4 Length: 760
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48745 (1-760)

QY 578 ACTCACTGTACGCTGTGGAACG 601
DB |||||||||||||||||||||

QY 137 ThrHisLeuTyraCysGlyThr 144
DB |||||||||||||||||||||

RESULT 48

T51787
hypothetical protein F28D10_90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51787
R:Deiseny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.; Le;
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25454
A:Accession: T51787
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-817
A:Cross-references: EMBL:AL391254
A:Experimental source: cultivar Columbia; BAC clone F28D10
C:Genetics:
A:Map position: 3
A:Introns: 88/3; 441/3
A:Note: F28D10_90

Alignment Scores:
Pred. No.: 34.1 Length: 817
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T51787 (1-817)

QY 1590 TTCCTAAGGAGACTTGGTATGATT 1613
DB |||||||||||||||||||||

Db 74 PheLeuArgArgLeuGlyWetille 81
|||||
RESULT 49
A86404
probable protein ATP-dependent DNA helicase RecQ [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86404
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-941 <STO>
A:Cross-references: GB:AE005172; NID:g1098928; PIDN:AG26068.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
Alignment Scores:
Pred. No.: 33.6 Length: 941
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 0.92%
Indels: 0
Gaps: 0
DB:
US-09-774-490-1 (1-2709) x A86404 (1-941)
QY 991 TCTAGCGTACGAGCTTTCCAGA 968
Db 98 SerSerValSerSerPheSerArg 105
|||||
RESULT 50
A33626
fibrinogen alpha chain - sea lamprey (fragment)
C:Species: Petromyzon marinus (sea lamprey)
C>Date: 09-Mar-1990 #sequence_revision 19-Feb-1999 #text_change 20-Aug-1999
C:Accession: A33626; A03120
R:Wang, Y.Z.; Patterson, J.; Gray, J.E.; Yu, C.; Cottrell, B.A.; Shimizu, A.; Graham, D.
Biochemistry 28, 9801-9806, 1989
A:Title: Complete sequence of the lamprey fibrinogen alpha-chain.
A:Reference number: A33626; MUID:90122867; PMID:2611265
A:Accession: A33626
A:Molecule type: DNA; mRNA; protein
A:Residues: 1-966 <WAN>
A:Cross-references: GB:M30123; NID:g213197; PIDN:AAA49263.1; PID:g213198
A:Note: residues 1-5 are not translated in Figure 2; parts of this sequence, including t
R:Cottrell, B.A.; Doolittle, R.F.
Biochim. Biophys. Acta 453, 426-438, 1976
A:Title: Amino acid sequences of lamprey fibrinopeptides A and B and characterization of
A:Reference number: A03120; MUID:77065679; PMID:999898
A:Accession: A03120
A:Molecule type: protein
A:Residues: 6-11 <COT>
C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology
C:Keywords: blood coagulation; glycoprotein; plasma
F:6-966/Product: fibrinogen alpha chain #status experimental <PRE>
F:36-164/Domain: fibrinogen disulfide ring homology <FDR>
F:920/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 33.5 Length: 966
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: Gaps: 0
US-09-774-490-1 (1-2709) x A33626 (1-966)
QY 153 ACAGGAGGAGAGACTAAGCAGCA 176
Db 949 ThrGlyArgLysThrLysAlaAla 956
|||||
RESULT 51
JC5928
semaphorin F precursor - human
C:Species: Homo sapiens (man)
C>Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C:Accession: JC5928
R:Simmons, A.D.; Puschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candid
A:Reference number: JC5928; MUID:98125554; PMID:9464278
A:Accession: JC5928
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1074 <SIM>
A:Cross-references: GB:U52840; NID:g2772583; PIDN:AAC09473.1; PID:g2772584
A:Experimental source: brain
C:Comment: This protein disrupts normal brain development and leads to some of the featu
C:Genetics:
A:Gene: semaf
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:50-533/Domain: semaphorin #status predicted <SEM>
F:840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F:971-993/Domain: transmembrane #status predicted <TMM>
Alignment Scores:
Pred. No.: 33.1 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: Gaps: 0
US-09-774-490-1 (1-2709) x JC5928 (1-1074)
QY 929 TACTTTTCTTCCTGAAATGCA 952
Db 230 TyrPhePhePheArgGluAsnAla 237
|||||
RESULT 52
F84487
probable ABC transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84487
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <STO>
A:Cross-references: GB:AE002093; NID:g5001458; PIDN:AAD37023.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g07680
A:Map position: 2
Alignment Scores:
Pred. No.: 32.9 Length: 1146

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 0 Gaps: 0

US-09-774-490-1 (1-2709) x F84487 (1-1146)

QY 1966 CTCTACACCATAGATGATTCTCTC 1943

Db 693 LeuTYrThrIleAspSerIleu 700

RESULT 53

A:Title: Hypothetical protein At2g25730 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: A84652
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84652
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2420 <STO>
A:Cross-references: GB:AE002093; NID:g4874311; PIDN:AA031373.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g25730
A:Map position: 2

Alignment Scores:

Pred. No.: 30.3 Length: 2420
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 0 Gaps: 0

US-09-774-490-1 (1-2709) x A84652 (1-2420)

QY 1996 GAAGTCGAGAGCGCTGGTCTA 2019

Db 1467 GluValAlaGluSerAlaGlyIleu 1474

RESULT 54

A:Title: ryanodine receptor type 1, skeletal muscle - human
N:Alternate names: calcium release channel protein
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A35041; I84622; S66630
R:Zorzato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meissner, G.;
J. Biol. Chem. 265, 2244-2256, 1990
A:Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release
A:Reference number: A35041; MUID:90130482; PMID:2298749
A:Accession: A35041
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-5032 <ZOR>
A:Cross-references: GB:J05200; NID:g337721; PIDN:AAA60294.1; PID:g337722
R:Otsu, K.; Phillips, M.S.; Khanna, V.K.; de Leon, S.; MacLennan, D.H.
Genomics 13, 835-837, 1992
A:Title: Refinement of diagnostic assays for a probable causal mutation for porcine and
A:Reference number: I46644; MUID:92347887; PMID:1639409
A:Accession: I84622
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 598-722 <RES>
A:Cross-references: GB:M91455; NID:g337723; PIDN:AAA60295.1; PID:g553643
R:Lynn, S.; Morgan, J.M.; Lamb, H.K.; Meissner, G.; Gillespie, J.I.

FEBS Lett. 372, 6-12, 1995

A:Title: Isolation and partial cloning of ryanodine-sensitive Ca(2+) release channel pro
A:Reference number: S66630; MUID:96032536; PMID:7556644

A:Accession: S66630

A:Molecule type: mRNA

A:Residues: 4690-4968 <LYN>

A:Experimental source: myometrial smooth muscle

C:Genetics:

A:Gene: GDB:RYR1

A:Cross-references: GDB:I20359; OMIM:180901

A:Map position: 19q13.1-19q13.1

A:introns: 642/2

C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
C:Keywords: calcium channel; homotetramer; phosphoprotein; skeletal muscle; transmembran
F:1788-1984/Domain: transcription initiation factor sigma region 1 homology <SR1>

Alignment Scores:

Pred. No.: 28 Length: 5032
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x A35041 (1-5032)

QY 64 GGGTCGACACCCCTCGACGTGGCA 41

Db 2268 GlySerThrProLeuAspValAla 2275

RESULT 55

I46646

A:Title: ryanodine receptor, skeletal muscle - pig

N:Alternate names: calcium release channel protein

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: I46646; I46645; S31395; I47133; S26624; A37105; I47212; S18135

R:Fujii, J.; Otsu, K.; Zorzato, F.; De Leon, S.; Khanna, V.K.; Weiler, J.E.; O'Brien, P.

Science 253, 448-451, 1991

A:Title: Identification of a mutation in porcine ryanodine receptor associated with mal

A:Reference number: I46645; MUID:91320118; PMID:1862346

A:Accession: I46646

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-5035 <FUJ>

A:Cross-references: GB:M91452; NID:g164647; PIDN:AAA31119.1; PID:g164648

A:Accession: I46645

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-614, 'C', 616-5035 <FU2>

A:Cross-references: GB:M91451; NID:g164645; PIDN:AAA31118.1; PID:g164646

R:Leeb, T.; Brem, G.; Brenig, B.

submitted to the EMBL Data Library, November 1992

A:Description: Genomic organization of porcine skeletal muscle ryanodine receptor gene c

A:Reference number: S31395

A:Accession: S31395

A:Molecule type: DNA

A:Residues: 1542-2643 <LEE>

A:Cross-references: EMBL:X69465

R:Leeb, T.; Schmoelzl, S.; Brem, G.; Brenig, B.

Genomics 18, 349-354, 1993

A:Title: Genomic organization of the porcine skeletal muscle ryanodine receptor (RYR1) g

A:Reference number: A48915; MUID:94117003; PMID:8288238

A:Contents: annotation

R:Harbilitz, I.; Kristensen, T.; Bosnes, M.; Kran, S.; Davies, W.

Anim. Genet. 23, 395-402, 1992

A:Title: DNA sequence of the skeletal muscle calcium release channel cDNA and verificati

A:Reference number: I47133; MUID:93036581; PMID:1329581

A:Accession: I47133

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 2-2091, 'A', 2093-3494, 'L', 3496-4163, 'S', 4165-4182, 'R', 4184-4411, 'W', 4413-4971

A:Cross-references: EMBL:X62880; NID:g1936; PIDN:CAA44674.1; PID:g1937
 R:Harbitz, I.; Kristensen, T.; Kran, T.; Davies, W.
 Submitted to the EMBL Data Library, August 1992
 A:Reference number: S26624
 A:Accession: S26624
 A:Molecule type: DNA
 A:Residues: 482-706 <HAW>
 A:Cross-references: EMBL:X68247
 R:Harbitz, I.; Chowdhary, B.; Thomsen, P.D.; Davies, W.; Kaufmann, U.; Kran, S.; Gustavsen
 Genomics 8, 243-248, 1990
 A:Title: Assignment of the porcine calcium release channel gene, a candidate for the mal
 A:Reference number: A37105; MUID:91065640; PMID:2174405
 A:Accession: A37105
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 4785-4971, 'R', 4973-5035 <HA2>
 A:Cross-references: GB:IM32501; NID:g164428; PIDN:AAA31022.1; PID:g164429
 R:Liedbetter, M.W.; Gruber, J.K.; Louis, C.F.; Mickelson, J.R.
 J. Biol. Chem. 269, 31544-31551, 1994
 A:Title: Tissue distribution of ryanodine receptor isoforms and alleles determined by re
 A:Reference number: A55660; MUID:95081095; PMID:7989322
 A:Accession: I47212
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 516-676 <LED>
 A:Cross-references: EMBL:U15965; NID:g562095; PIDN:AAA60467.1; PID:g562096
 C:Genetics:
 A:Gene: RYR1
 A:Introns: 327/1; 559/1; 598/3; 643/2; 1570/3; 1646/2; 1850/3; 1939/3; 2006/3; 2044/1; 2
 A>Note: the list of introns may be incomplete
 A:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
 C:Keywords: calcium channel
 Alignment Scores:
 Pred. No.: 28 Length: 5035
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x I46646 (1-5035)
 QY 64 GGGTCGACACCCCTCGACGTGGCA 41
 Db 2270 GlySerThrProLeuAspValAla 2277
 RESULT 56
 B35041
 ryanodine receptor, skeletal muscle - rabbit
 N:Alternate names: calcium-release channel protein; junctional channel complex
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 27-Jul-1990 #sequence revision 10-Mar-1994 #text_change 20-Aug-1999
 A:Accession: S04654; B35041; A36181; S53794; S32504
 R:Takeshima, H.; Nishimura, S.; Matsumoto, T.; Ishida, H.; Kangawa, K.; Minamino, N.; Ma
 Nature 339, 439-445, 1989
 A:Title: Primary structure and expression from complementary DNA of skeletal muscle ryan
 A:Reference number: S04654; MUID:89262082; PMID:2725677
 A:Accession: S04654
 A:Molecule type: mRNA
 A:Residues: 1-5037 <TAK>
 A:Cross-references: EMBL:X15750; NID:g1709; PIDN:CAA33762.1; PID:g1710
 A>Note: part of this sequence was confirmed by protein sequencing
 R:Zorzato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meisner, G.;
 J. Biol. Chem. 265, 2244-2256, 1990
 A:Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release
 A:Reference number: A35041; MUID:90130482; PMID:2298749
 A:Accession: B35041
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-66, 'Y', 68-2014, 'D', 2016-3246, 'E', 3248-3480, 3486-4497, 'LE', 4498-4521, 'Q', 45
 4748-4758, 'N', 4760-5037 <ZOR>
 R:Marks, A.R.; Tempst, P.; Hwang, K.S.; Taubman, M.B.; Inui, M.; Chadwick, C.; Fleischer

Proc. Natl. Acad. Sci. U.S.A. 86, 8683-8687, 1989
 A:Title: Molecular cloning and characterization of the ryanodine receptor/junctional cha
 A:Reference number: A36181; MUID:90046857; PMID:2813419
 A:Accession: A36181
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA; protein
 A:Residues: 'X', 1223-1235, 'XX', 1238-1251; 1334-1348; 1566-1569, 'X', 1571, 'X', 1573; 1597-1604
 -4679, 'X', 4681-4689, 'X', 4691-4693, 'X', 4695, 'X', 4697-4700 <MAR>
 A>Note: the proteolytic fragments sequenced here from the junctional channel complex as
 R:Varsanyi, M.; Meyer, H.E.
 Biol. Chem. Hoppe-Seyler 376, 45-49, 1995
 A:Title: Sarcoplasmic reticular Ca(2+) release channel is phosphorylated at serine 2843
 A:Reference number: S53794; MUID:95336639; PMID:7612188
 A:Accession: S53794
 A:Molecule type: protein
 A:Residues: 2841-2852 <VAR>
 R:Takeshima, H.; Nishimura, S.; Nishi, M.; Ikeda, M.; Sugimoto, T.
 FEBS Lett. 322, 105-110, 1993
 A:Title: A brain-specific transcript from the 3'-terminal region of the skeletal muscle
 A:Reference number: S32504; MUID:93245969; PMID:8097730
 A:Accession: S32504
 A:Molecule type: mRNA
 A:Residues: 4163-5037 <TAW>
 C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
 C:Keywords: calcium channel; glycoprotein; phosphoprotein; skeletal muscle; transmembran
 P:695-704/Region: adenine nucleotide binding
 P:841-954, 955-1068/Region: 114-residue repeats
 P:1344-1359, 1371-1386/Region: 16-residue repeats
 P:1789-1985/Domain: adenine nucleotide binding
 P:2370-2375/Region: adenine nucleotide binding
 P:2725-2844, 2845-2958/Region: 120-residue repeats
 P:4564-4580/Domain: transmembrane #status predicted <TM1>
 P:4641-4664/Domain: transmembrane #status predicted <TM2>
 P:4836-4859/Domain: transmembrane #status predicted <TM3>
 P:4918-4937/Domain: transmembrane #status predicted <TM4>
 P:128, 286/Binding site: phosphate (Thr) (covalent) by CAMP- and calmodulin-dependent ki
 P:2843/Binding site: phosphate (Ser) (covalent) #status experimental
 P:4864/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Alignment Scores:
 Pred. No.: 28 Length: 5037
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 2 Gaps: 0
 US-09-774-490-1 (1-2709) x B35041 (1-5037)
 QY 64 GGGTCGACACCCCTCGACGTGGCA 41
 Db 2269 GlySerThrProLeuAspValAla 2276
 RESULT 57
 T30192
 Probable peptide synthetase - Aureobasidium pullulans
 C:Species: Aureobasidium pullulans
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C:Accession: T30192
 R:Peery, R.B.; Thornnewell, S.J.; Tobin, M.B.; Skatrud, P.L.
 submitted to the EMBL Data Library, January 1997
 A:Description: Discovery of an MDR-like gene adjacent to a peptide synthetase in Aureoba
 A:Reference number: Z20767
 A:Accession: T30192
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-10797 <PEE>
 A:Cross-references: EMBL:U85909; NID:g4099310; PID:g4099313; PIDN:AAD00581.1
 C:Genetics:
 A:Introns: 2078/3; 4142/3; 6075/3; 7963/3; 7985/2
 C:Superfamily: acyl carrier protein homology
 C:Keywords: carrier protein
 P:1618-1688/Domain: acyl carrier protein homology <ACP1>

F:3682-3752/Domain: acyl carrier protein homology <ACP2>
 F:5635-5685/Domain: acyl carrier protein homology <ACP3>
 F:7503-7573/Domain: acyl carrier protein homology <ACP4>
 F:9683-9752/Domain: acyl carrier protein homology <ACP5>

Alignment Scores:
 Pred. No.: 25.7 Length: 10797
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T30192 (1-10797)

QY 645 ATCTGAGGACATATTTTAAAGC 668

Db 1781 lleleuArgThrilePheleuser 1788

RESULT 58

S09700

phycobiliprotein 18.3 beta chain - *Synechococcus* sp. (PCC 6301) (fragment)

C:Species: *Synechococcus* sp.

A:Variety: PCC 6301

C>Date: 21-Nov-1993 #sequence_revision 01-Mar-1996 #text_change 09-May-1997

R:Accession: S09700

R:Lundell, D.J.; Glazer, A.N.

J. Biol. Chem. 258, 894-901, 1983

A:Title: Molecular architecture of a light-harvesting antenna. Structure of the 18 S core

A:Reference number: S09700; PMID:83108897; PMID:6401720

A:Accession: S09700

A:Molecule type: protein

A:Residues: 1-16 <LUN>

A:Experimental source: PCC 6301

Alignment Scores:

Pred. No.: 630 Length: 16
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S09700 (1-16)

QY 738 AGGATGCTGTGACGAGCTTA 718

Db 2 ArgAspAlaValSerSerleu 8

RESULT 59

B40256

interleukin-7 receptor, soluble form - human (fragment)

C:Species: *Homo sapiens* (man)

C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 23-May-1997

R:Accession: B40256

R:Pfeiman, C.M.; Gimpel, S.D.; Park, L.S.; Hazada, H.; Taniguchi, T.; Ziegler, S.F.

Mol. Cell. Biol. 11, 3052-3059, 1991

A:Title: Organization of the murine and human interleukin-7 receptor genes: two mRNAs ge

A:Reference number: A40256; PMID:91246172; PMID:2038316

A:Accession: B40256

A:Molecule type: DNA

A:Residues: 1-22 <PLE>

A>Note: the authors translated the codon GAA for residue 20 as Gln

C:Superfamily: interleukin-7 receptor; fibronectin type III repeat homology

C:Keywords: cytokine receptor

Alignment Scores:

Pred. No.: 608 Length: 22
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B40256 (1-22)

QY 918 TCCTCAGGATGTCTACTCTCT 898

Db 4 SerSerGlyLeuSerLeuSer 10

RESULT 60

B82805

hypothetical protein XF0448 [imported] - *Xylella fastidiosa* (strain 9a5c)

C:Species: *Xylella fastidiosa*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82805

R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; PMID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: B82805

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-35 <SIM>

A:Cross-references: GB:AE003895; GB:AE003849; NID:9105283; PIDN:AAF83258.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaas

A:Authors: da Silva, A.C.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0448

Alignment Scores:

Pred. No.: 578 Length: 35
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B82805 (1-35)

QY 402 CTACTCCGTCCTCATCCAA 382

Db 15 LeuLeuArgSerSerSerLys 21

RESULT 61

S03674

photosystem I protein psaj - *Euglena gracilis* chloroplast

C:Species: chloroplast *Euglena gracilis*

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-Aug-1999

C:Accession: S03674; S26069

R:Manzara, T.; Hallick, R.B.

Nucleic Acids Res. 16, 9866, 1988

A:Title: Nucleotide sequence of the *Euglena gracilis* chloroplast genes for serine and pr

A:Reference number: S03674; PMID:89041581; PMID:3141903

A:Accession: S03674

A:Molecule type: DNA

A:Residues: 1-37 <MAN>

A:Cross-references: EMBL:M18672; NID:G336891; PIDN:AAA84231.1; PID:G896264

C:Genetics:

A:Gene: psaj

A:Genome: chloroplast

C:Superfamily: photosystem I protein psaj

C:Keywords: chloroplast; photosystem I; transmembrane protein

Alignment Scores:
Pred. No.: 575 Length: 37
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S03674 (1-37)

QY 725 CTGACAGATCCCTTTTAATA 745

DB 21 LeuThrAlaSerLeuLeu 27

RESULT 62

S05490

alpha-amylase (EC 3.2.1.1) 2.46 precursor - wheat (fragment)

C:Species: Triticum aestivum (common wheat)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 06-Dec-1996

C:Accession: S05490

R:Huttlly, A.K.; Martienssen, R.A.; Baulcombe, D.C.

Mol. Gen. Genet. 214, 232-240, 1988

A:Title: Sequence heterogeneity and differential expression of the alpha--Amy-2 gene fam

A:Reference number: S05486; MUID:89181522; PMID:2467183

A:Accession: S05490

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-51 <HT>

A:Cross-references: EMBL:X13579

C:Genetics:

A:Gene: amy2

A:Map position: 7A

A:Introns: 14/1

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycohen/starch degradation

C:Superfamily: wheat alpha-amylase; alpha-amylase core homology

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Alignment Scores:

Pred. No.: 555 Length: 51
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S05490 (1-51)

QY 179 CTTTCTGCTTTAGCTTCCT 159

DB 8 LeuCysCysPheSerLeuPro 14

RESULT 63

D42194

pepH protein - Synecococcus sp. (PCC 7002) (fragment)

C:Species: Synecococcus sp.

C:Date: 04-Mar-1993 #sequence_revision 02-Aug-1994 #text_change 26-Aug-1999

C:Accession: D42194

R:Schluchter, W.M.; Bryant, D.A.

Biochemistry 31, 3092-3102, 1992

A:Title: Molecular characterization of ferredoxin-NADP+ oxidoreductase in cyanobacteria:

A:Reference number: A42194; MUID:92207922; PMID:1554697

A:Accession: D42194

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-63 <SCH>

A:Cross-references: GB:M86234; GB:J05366; NID:g154535; PIDN:AAA27325.1; PID:g154538

A:Experimental source: PCC 7002

A:Note: sequence extracted from NCBI backbone (NCBIN:91780, NCBIP:91792)

C:Superfamily: ompR protein; response regulator homology

F:21-63/Domain: response regulator homology (fragment) <RRH>

Alignment Scores:
Pred. No.: 542 Length: 63
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D42194 (1-63)

QY 110 GCTGACATGGGAGACAGGC 90

DB 48 AlaAspAsnGlyArgThrGly 54

RESULT 64

S33581

hypothetical protein 3 - Azorhizobium caulinodans

C:Species: Azorhizobium caulinodans

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-May-2000

C:Accession: S33581

R:Stigter, J.; Schneider, M.; de Bruijn, P.J.

Mol. Plant Microbe Interact. 6, 238-252, 1993

A:Title: Azorhizobium caulinodans nitrogen fixation (nif/fix) gene regulation: mutagenes

A:Reference number: S33579; MUID:93229807; PMID:8471796

A:Accession: S33581

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-68 <STI>

A:Cross-references: EMBL:X69959; NID:g311387; PIDN:CAA49583.1; PID:g311390

Alignment Scores:

Pred. No.: 537 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S33581 (1-68)

QY 1719 CTGGGGTTGCCAGCTCCCTT 1739

DB 24 LeuGlyLeuProSerSerLeu 30

RESULT 65

A83261

hypothetical protein PA3085 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83261

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83261

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-87 <STO>

A:Cross-references: GB:AE004732; GB:AE004091; NID:g9949186; PIDN:AAG06473.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3085

Alignment Scores:

Pred. No.: 523 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0

DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x AB3261 (1-87)
QY 284 CATTGTTCTCCATTCATTCGAT 264
DB 67 HiscysSerSerHisSerAsp 73
RESULT 66
T25449
hypoethetical protein B0412.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000
C:Accession: T25449
R:Bentley, D.
A:Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid B0412.
A:Reference number: Z20037
A:Accession: T25449
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88 <SEN>
A:Cross-references: EMBL:U0953; PIDN:AAB52557.1; GSPDB:GN00021; CESP:B0412.4
A:Experimental source: strain Bristol N2; clone B0412
C:Genetics:
A:Gene: CESP:B0412.4
A:Map position: 3
A:Introns: 4/3; 28/2; 61/3
C:Superfamily: Escherichia coli ribosomal protein S14
Alignment Scores:
Pred. No.: 522 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T25449 (1-88)
QY 2406 CTTGCGCAGCTGTTTCGG 2386
DB 45 LeuCySArg-gcysPheArg 51
RESULT 67
T30626
hypoethetical protein 24L - Molluscum contagiosum virus 1
N:Alternate names: MC024L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-May-2000
C:Accession: T30626
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30626
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88 <SEN>
A:Cross-references: EMBL:U60315; PIDN:AAC55152.1
C:Genetics:
A:Note: MC024L
Alignment Scores:
Pred. No.: 522 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T30626 (1-88)

QY 40 ATGCCACGTCGAGGGTGTCTG 60
DB 38 MetProArgGlyValser 44
RESULT 68
I68530
hemoglobin alpha chain (clone alphaG-28) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 19-May-2000
C:Accession: I68530
R:Lam, V.M.; Gu, Y.L.; Au, D.M.; Wong, W.M.; Ma, C.W.; Cheng, L.Y.
Hemoglobin 17, 363-371, 1993
A:Title: Two new rat alpha-globin sequences as identified by the conserved region PCR.
A:Reference number: I54239; MUID:94042225; PMID:8226096
A:Accession: I68530
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <RES>
A:Cross-references: GB:S66658; NID:9439804
C:Genetics:
A:Introns: 32/2
C:Superfamily: globin; globin homology
C:Keywords: blood; oxygen carrier
F:59/Binding site: oxygen (His) (distal axial ligand) #status predicted
Alignment Scores:
Pred. No.: 522 Length: 89
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I68530 (1-89)
QY 1464 ACTATTGGCGATTGTTTCATA 1444
DB 29 ThrileGlyArgLeuPheille 35
RESULT 69
AF2068
hypoethetical protein asl2100 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AF2068
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ara
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2068
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA873799.1; PID:G17131191; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asl2100
Alignment Scores:
Pred. No.: 520 Length: 91
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x AF2068 (1-91)
QY 1969 ATTCTTACACCATGATGAT 1949
DB 63 IleLeuTyThrIleAspAsp 69

Search completed: August 3, 2003, 10:34:52
Job time : 186.5 secs

C:Accession: D70408
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: D70408
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-102 <AQF>
A:Cross-references: GB:AE000731; NID:g2983691; PIDN:AAC07266.1; PID:g2983702; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_1254
C:Superfamily: yajC protein

Alignment Scores:
Pred. No.: 514 Length: 102
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D70408 (1-102)

QY 1404 AAGTTTAAATCATCATCAGGA 1384

Db 51 LysValIleThrSerGly 57

RESULT 75

IPXL2

insulin II precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: B33847; S13537
R:Shuldiner, A.R.; Phillips, S.; Roberts Jr., C.T.; LeRoith, D.; Roth, J.
J. Biol. Chem. 264, 9428-9432, 1989
A:Title: Xenopus laevis contains two nonallelic preproinsulin genes. cDNA cloning and ev
A:Reference number: A33847; MUID:89255444; PMID:2722842
A:Accession: B33847
A:Molecule type: mRNA
A:Residues: 1-106 <SHU1>
A:Cross-references: GB:M24442; GB:J04796; NID:g214534; PIDN:AAA49887.1; PID:g214535
R:Shuldiner, A.R.; Bennett, C.; Robinson, E.A.; Roth, J.
Endocrinology 125, 469-477, 1989
A:Title: Isolation and characterization of two different insulins from an amphibian, Xen
A:Reference number: S07199; MUID:89289601; PMID:2661211
A:Accession: S13537

A:Molecule type: protein
A:Residues: 24-53; 86-106 <SHU2>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-53/Domain: insulin chain B #status experimental <BCH>
F:24-53; 86-106/Product: insulin #status experimental <MAT>
F:56-83/Domain: connecting peptide #status predicted <CEP>
F:86-106/Domain: insulin chain A #status experimental <ACH>
F:30-92, 42-105, 91-96/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.: 512 Length: 106
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x IPXL2 (1-106)

QY 1977 GTGCTACTATCTCTACACCA 1957

Db 13 ValLeuLeuPheSerThrPro 19

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 15:53:26 ; Search time 597 Seconds
(without alignments)
9361.255 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatctttattttatcgatg.....aggcttttttctctaataacc 2709

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2709	100.0	2709	9	US-09-774-490-1
2	2508.4	92.6	2530	14	Sequence 1, Appli
3	2508.4	92.6	2530	14	Sequence 283, App
4	718.6	26.5	3871	11	Sequence 9, Appli
5	718.6	26.5	3871	11	Sequence 309, App
6	718.6	26.5	3871	13	Sequence 347, App
7	718.6	26.5	3871	14	Sequence 347, App
8	718.6	26.5	3871	14	Sequence 347, App
9	718.6	26.5	3871	14	Sequence 347, App
10	718.6	26.5	3871	14	Sequence 347, App
11	718.6	26.5	3871	14	Sequence 347, App
12	718.6	26.5	3871	14	Sequence 347, App
13	718.6	26.5	3871	14	Sequence 347, App
14	718.6	26.5	3871	14	Sequence 347, App
15	718.6	26.5	3871	14	Sequence 347, App
16	718.6	26.5	3871	14	Sequence 347, App

17	718.6	26.5	3871	14	US-10-173-700-347	Sequence 347, App
18	718.6	26.5	3871	14	US-10-174-572-347	Sequence 347, App
19	718.6	26.5	3871	14	US-10-174-579-347	Sequence 347, App
20	718.6	26.5	3871	14	US-10-174-582-347	Sequence 347, App
21	718.6	26.5	3871	14	US-10-174-588-347	Sequence 347, App
22	718.6	26.5	3871	14	US-10-175-739-347	Sequence 347, App
23	718.6	26.5	3871	14	US-10-175-740-347	Sequence 347, App
24	718.6	26.5	3871	14	US-10-175-743-347	Sequence 347, App
25	718.6	26.5	3871	14	US-10-176-488-347	Sequence 347, App
26	718.6	26.5	3871	14	US-10-176-492-347	Sequence 347, App
27	718.6	26.5	3871	14	US-10-176-747-347	Sequence 347, App
28	718.6	26.5	3871	14	US-10-176-750-347	Sequence 347, App
29	718.6	26.5	3871	14	US-10-176-985-347	Sequence 347, App
30	718.6	26.5	3871	14	US-10-176-987-347	Sequence 347, App
31	718.6	26.5	3871	14	US-10-176-992-347	Sequence 347, App
32	718.6	26.5	3871	14	US-10-176-993-347	Sequence 347, App
33	718.6	26.5	3871	14	US-10-184-658-347	Sequence 347, App
34	718.6	26.5	3871	14	US-10-176-991-347	Sequence 347, App
35	718.6	26.5	3871	14	US-10-173-695-347	Sequence 347, App
36	718.6	26.5	3871	14	US-10-173-697-347	Sequence 347, App
37	718.6	26.5	3871	14	US-10-173-705-347	Sequence 347, App
38	718.6	26.5	3871	14	US-10-174-576-347	Sequence 347, App
39	718.6	26.5	3871	14	US-10-174-585-347	Sequence 347, App
40	718.6	26.5	3871	14	US-10-174-586-347	Sequence 347, App
41	718.6	26.5	3871	14	US-10-175-747-347	Sequence 347, App
42	718.6	26.5	3871	14	US-10-176-481-347	Sequence 347, App
43	718.6	26.5	3871	14	US-10-176-485-347	Sequence 347, App
44	718.6	26.5	3871	14	US-10-176-487-347	Sequence 347, App
45	718.6	26.5	3871	14	US-10-176-493-347	Sequence 347, App

ALIGNMENTS

RESULT 1

US-09-774-490-1
; Sequence 1, Application US/09774490
; Patent No. US20010034332A1
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1

Query Match	100.0%	Score 2709;	DB 9;	Length 2709;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2709;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AATCTTTTATTTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGCG	60	
DB	1	AATCTTTTATTTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGCG	60	
QY	61	ACCCACGGTCGGGAGTAGTTGAGCTCGCCTGTTCCTCCATTCTCAGCCAGTCTATT	120	
DB	61	ACCCACGGTCGGGAGTAGTTGAGCTCGCCTGTTCCTCCATTCTCAGCCAGTCTATT	120	
QY	121	CCAGATTGTTTGAACCTTCTTGGCCGCAATACAGGAAGGAAGACTAAACGACAAAGG	180	
DB	121	CCAGATTGTTTGAACCTTCTTGGCCGCAATACAGGAAGGAAGACTAAACGACAAAGG	180	
QY	181	GACCTACAGCTGTCAGCATGGGCTGGTGGTAACTAGGATTCGTCTCTTTCTTGGGGAGT	240	
DB	181	GACCTACAGCTGTCAGCATGGGCTGGTGGTAACTAGGATTCGTCTCTTTCTTGGGGAGT	240	

Db 2401 GCAAAGGCCAGACATACCCAGGGAACAGTAAACAAATGGAAGCCTTACAAAGAAATAA 2460
QY 2461 GAAAGGTAGAAACAGGAGGACCCAGCAATTTGAGAGGGGACCCAGGAGTGTCTGAGCTGC 2520
Db 2461 GAAAGGTAGAAACAGGAGGACCCAGCAATTTGAGAGGGGACCCAGGAGTGTCTGAGCTGC 2520
QY 2521 ATTACTCTAGAAACCTCAACAGTAGAGAACTTGCCTAGACAATACTGGGAAAAACAAA 2580
Db 2521 ATTACTCTAGAAACCTCAACAGTAGAGAACTTGCCTAGACAATACTGGGAAAAACAAA 2580
QY 2581 TCGCAATATACATGAACCTTTTTCATGGCAATATGAGGATGTTTACAATGGTGGGAAATTC 2640
Db 2581 TCGCAATATACATGAACCTTTTTCATGGCAATATGAGGATGTTTACAATGGTGGGAAATTC 2640
QY 2641 AGCTGAGTTCCACCAATTATAAATAAATCCATGAGTAACCTTCCCTAATAGGCTTTTTTT 2700
Db 2641 AGCTGAGTTCCACCAATTATAAATAAATCCATGAGTAACCTTCCCTAATAGGCTTTTTTT 2700
QY 2701 CCTAATACC 2709
Db 2701 CCTAATACC 2709

RESULT 2

US-10-097-340-283
; Sequence 283, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHWANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-283

Query Match 92.6%; Score 2508.4; DB 14; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 193 CTGACAGCATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 252
Db 9 CTGACAGCATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATATCTCAAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATATCTCAAAAGA 128
QY 313 AATGTTGGAATCCCAACATGTGATCCTTTCAATGGCTTGGCCACACAGCTCCAGTATCA 372
Db 129 AATGTTGGAATCCCAACATGTGATCCTTTCAATGGCTTGGCCACACAGCTCCAGTATCA 188
QY 373 TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGAGCAAAAGGATCACATATT 432
Db 189 TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGAGCAAAAGGATCACATATT 248
QY 433 TTCAATCGACCTGGTTAATATCAAGGAATTTTCAAAAGATTTGTGCGCCAGTATCTTACAC 492
Db 249 TTCAATCGACCTGGTTAATATCAAGGAATTTTCAAAAGATTTGTGCGCCAGTATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTCTGAAAAGATGTCTAATTTGAT 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTCTGAAAAGATGTCTAATTTGAT 368
QY 553 CAAGGTACTTAAAGGCATATAATCAGACTCATTGTGACCCCTGTGGAACGGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAAGGCATATAATCAGACTCATTGTGACCCCTGTGGAACGGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAAATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAAATATTTTAAAGCTGA 488
QY 673 GAACTCACAATTTGAAAACGGCCGTGGGAAAGTCCATATGACCCTAAGTCTGCACAGC 732
Db 489 GAACTCACAATTTGAAAACGGCCGTGGGAAAGTCCATATGACCCTAAGTCTGCACAGC 548
QY 733 ATCCCTTTTAAATAGATGAGAAATATATCTCTGGAATCTCAGCTGATTTTATGGGCGGAGA 792
Db 549 ATCCCTTTTAAATAGATGAGAAATATATCTCTGGAATCTCAGCTGATTTTATGGGCGGAGA 608
QY 793 CTTTGCTATCTTCCGAATCTTTGGSCACACCACCACCAATCAGGACAGAGCATGATTC 852
Db 609 CTTTGCTATCTTCCGAATCTTTGGSCACACCACCACCAATCAGGACAGAGCATGATTC 668
QY 853 CAGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCC 912
Db 659 CAGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCC 728
QY 913 TGAAGATGACAAAAGTATATCTTTCTCCGTGAAAATCAATAGATGAGAACTCTCTGG 972
Db 729 TGAAGATGACAAAAGTATATCTTTCTCCGTGAAAATCAATAGATGAGAACTCTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACCTTTGGAGGSCACAGAAG 1032
Db 789 AAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACCTTTGGAGGSCACAGAAG 848
QY 1033 TCTGGTGAATAAATGGACAACATTCCTCAAAAGTCTGTGATTTGCTCAGTGGCCAGGTCC 1092
Db 849 TCTGGTGAATAAATGGACAACATTCCTCAAAAGTCTGTGATTTGCTCAGTGGCCAGGTCC 908
QY 1093 AAATGGCATTGACACTCATTTTGTATGAACTCAGAGTGTATCTCTAATGAATTTTAAAGA 1152
Db 909 AAATGGCATTGACACTCATTTTGTATGAACTCAGAGTGTATCTCTAATGAATTTTAAAGA 968
QY 1153 TCCTAAAAATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC 1212
Db 969 TCCTAAAAATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC 1028
QY 1213 AGCCGTGTATGTATAGCATGTATGTGAGAGGGTGTCTTGGTTCATATGCCCCA 1272

Db 1029 AGCCGCTGTGTATGTATAGCATGAGTGTGAGAGGGGTGTTCTTCTGTGTCCTATATGCCCA 1088
Qy 1273 CAGGATGACCAACCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTCTATCCAGGCC 1332
Db 1089 CAGGATGACCAACCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTCTATCCAGGCC 1148
Qy 1333 AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACCTTCTCTGATGA 1392
Db 1149 AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGTACTCTCAAAAGGACCTTCTCTGATGA 1208
Qy 1393 TGTATTAACCTTTGCAAGAGTCATCCAGGCATGTACAAATCAAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCATCCAGGCATGTACAAATCAAGTGTTCCTATGAACAA 1268
Qy 1453 TCGCCCAATAGTATGATCAAAACCGATGTAATTTATCAAAATTTGCTGTAGACCG 1512
Db 1269 TCGCCCAATAGTATGATCAAAACCGATGTAATTTATCAAAATTTGCTGTAGACCG 1328
Qy 1513 AGTGGATGCAAGAGTGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGGATGCAAGAGTGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1388
Qy 1573 TCTTAAAGTAGTTCATTTCTTAAGGAGCTTGGTATGATTTAGAGAGGTTCTGCTGGA 1632
Db 1389 TCTTAAAGTAGTTCATTTCTTAAGGAGCTTGGTATGATTTAGAGAGGTTCTGCTGGA 1448
Qy 1633 AGAATGACAGTTCCTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACCTAAGCA 1692
Db 1449 AGAATGACAGTTCCTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACCTAAGCA 1508
Qy 1693 GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACGGTGTGA 1752
Db 1509 GCAACAACTATATTTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACGGTGTGA 1568
Qy 1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGTGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGTGGA 1628
Qy 1813 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGACAGATAT 1872
Db 1629 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGACGACAGATAT 1688
Qy 1873 AAGAAATGGAGACCACTGACTCAGTGTTCAGACTTTACACCATGATATCAACATGGCCA 1932
Db 1689 AAGAAATGGAGACCACTGACTCAGTGTTCAGACTTTACACCATGATATCAACATGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGATGCG 1992
Db 1749 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGATGCG 1808
Qy 1993 TCCGAAGTCGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGCAAAATGAAGACGAAA 2052
Db 1809 TCCGAAGTCGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGCAAAATGAAGACGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGCTACGTAG 1928
Qy 2113 TCTACAAAGAGAGATTAGGCAATTAACCTCTGCCATGCGGTGGAAACATGGGTTCAATCA 2172
Db 1929 TCTACAAAGAGAGATTAGGCAATTAACCTCTGCCATGCGGTGGAAACATGGGTTCAATCA 1988
Qy 2173 AACTCTTCTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCATTTGGGAAGACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAAACCTCGAAGTCAATTCACACAGAGCATTTGGGAAGACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGGTCTGGTACAGAGACTTCAATGAGCTTCATCAACACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGGTACAGAGACTTCAATGAGCTTCATCAACACCCCAATCTCAACAGAT 2168

Qy 2353 GGATGAGTTCCTGTGAACCAAGTTTGGAAAAGGACCGAAAAACAACTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTTCCTGTGAACCAAGTTTGGAAAAGGACCGAAAAACAACTCGGCAAGGCCAGG 2228
Qy 2413 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAATAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAATAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGACCCACGAATTTGAGAGGGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCACGAATTTGAGAGGGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2348
Qy 2533 AACCTCAAAACAGTAGAAAACCTTGCCTAGACAATAACTGGAATAACAAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAGTAGAAAACCTTGCCTAGACAATAACTGGAATAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGGCATTATGTGGATGTTTACAATGTTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGGCATTATGTGGATGTTTACAATGTTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTC 2702
Db 2469 CCAATTATAAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTC 2518

RESULT 3

US-10-262-538-9
; Sequence 9, Application US/10262538
; Publication No. US2003011324A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)..(2331)
US-10-262-538-9

Query Match 92.6%; Score 2508.4; DB 14; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 193 CTGCAGCATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTACAGC 68
Qy 253 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCAAGGCTGAAATTTCTTACAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCAAGGCTGAAATTTCTTACAAAGA 128
Qy 313 AATGTTGGAATCCAAACAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAAACAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
Qy 373 TACCTTCCTTTTGGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGAGGATCACATATT 432
Db 189 TACCTTCCTTTTGGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGAGGATCACATATT 248
Qy 433 TTCAATTCACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGATCTTTACAC 492
Db 249 TTCAATTCACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGATCTTTACAC 308
Qy 493 CAGAAGAGATGAATGCAAGTGGCTGGAAAAGACATCTCTGAAAGAAATGTGCTAAATTCAT 552

RESULT 4
US-09-946-374-309
Sequence 309, Application US/09946374
Publication No. US20030073129A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2830PIC1
CURRENT APPLICATION NUMBER: US/09/946,374
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099602
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099642
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099754
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099808
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100388
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100584
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100661
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100664
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29

```

; PRIOR APPLICATION NUMBER: 60/102484
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102487
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102571
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102684
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102687
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102965
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/103258
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103314
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103315
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103328
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103395
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103401
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103633
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807

Query Match          26.5%; Score 718.6; DB 11; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

Qy 269 AATGGGAAGAACATGTGCGAAGGCTGAAATATCTACAAAGAAATGTGGAAATCCAAAC 328
Db 215 ACTTTGAAGCAAAATATTCRAAGACTCAAGCTAACTACAAAGACTTGTCTCTTCAAT 274
Qy 329 AATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGCTTATATCATCTCTCTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACTCTTCTCTTAGAT 334
Qy 389 GAGGAACGAGTAGCTGTATGTTCGGCAAGAGTACATATTTTCATTTCGACCTGGTT 448
Db 335 GAGGAAGAGGAGGCTGCTCTTGGGAGCCAAAGACCACATCTTTCTACTCAGTCTGGTT 394

449 AA----TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
395 GACTTTAAACAAAAATTTTAAAGAGATTTATTGGCCCTGTGCAAGGAACGGGTGGAATTA 454
506 TGCAAGTGGGCTGGAAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTACTTAAAG 565
455 TGTAAATTTAGCTGGGAAAGATGCCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514
566 GCATATAATCAGACTCACTGTAGCCCTGTGGAACGGGGCTTTTCATCCCAATTTTGCACC 625
515 CCCTATAACAAACTCACAATATATGTGTGGAACTGGAGCAATTTTCATCCAAATATGTGGG 574
626 TACATTGAAATTTGGACATCATCTCTGAGGCAATATTTTTAAGCTGGAGAACTCACATTTT 685
575 TATATTGATCTTGGAGTCTACAAGGAGGATATATATTCAAACCTAGACACACATAATTTG 634
686 GAAAACGGCCGTGGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
635 GAGTCTGGCAGACTGAAATGTCTTTCGATCCTCAGCAGCCCTTTTGTCTCAGTAAATGACA 694
746 GATGAGAAATTTATCTCTGGAACTGCAGCTGATTTTATGGGCGGAGACTTTTGTCTATCTTC 805
695 GATGATACCTCTACTCTGGAAACAGCTTCTGATTTCTTGGCAAGAGACTACTGCATTTCACT 754
806 CGAACTCTTTGGGC-----ACCACCAACCAATCAGGACAGAGCAGCATGATTTCCAGG 856
755 CGATCCCTTTGGGCCCTACTCATGACCACCACTACATCAGAACTGCACATTTTCAGAGCCTAC 814
857 TGGCTCAATGATCCAAAGTTCAATTAGTCCCACTCATCTCAGAGAGTGCACATCTCTGNA 916
815 TGGCTCAATGAGGCAAAATTTTATTGGAACTTTCTTCATACCAGACACCTACAATCCAGAT 874
917 GATGACAAAGTATATCTTTTCTTCGTAAGAAATGCAATAGATGGAGAACACTCTCTGAAAA 976
875 GATGATAAAATATATTTCTTCTTTCGTAATCATCTCAAGAAAGCAGTACCTCCGATAAA 934
977 GCTACTACGCTAGATAAGTGCAGATATGCAAGAAATGATTTTGGAGGCAACAGAGTCTG 1036
935 ACCATCTCTTCTCGAGTTGGAAGAGTTTGTGAAGATGATGTAGGAGGCAACAGCGACCTG 994
1037 GTGAATAAATGGACAACTTCTCAAGCTCGTGTGATTTGCTCAGTCCAGGTCCTCAAT 1096
995 ATAAACCAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTCTGGAAGTGA 1054
1097 GGCATTGACACTCAATTTTGTAGAACTGCAGAGATGATTTCTTAAGAACTTTTAAAGATCCT 1156
1055 GGGCAGATATCTTACTTTGATGAGCTTCAAGATATTTTATTTACTCCCAAGAGATGAA 1114
1157 AAAAATCCAGTTGATATGAGTGTTCACGACTTTCAGTAACTTTTCAAGGGGATCAGCC 1216
1115 AGAAATCCTGTAGTATATGAGTCTTTTACTACAACCACTCCATCTTCAAGGGCTCTGCT 1174
1217 GTGTGATATGATAGATGATGATGAGAGGGTGTCTTGTGCTCATATGCCACAGG 1276
1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTGTTTTAAATGGTCTCATATGCTCATAG 1234
1277 GATGACCCCAACTATCAATGGTGCCTTTATCAAGAAAGAGTCCCTCTATCCACGGCCAGA 1336
1235 GAAAGTGCAGACCAATCGTTGGGTGCAGTATGATGGGAGAAATTCCTTATCCAGGGCTGGT 1294
1337 ACTTGTCCCAAGCAAAACATTTG----GTGGTTTTGACTCTTACAAAGGAGCTTTCTCTGATGAT 1393
1295 ACATGTCCAAGCAAAACCTATGACCCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
1394 GTTATAACCTTTGCAAGAGTGCATCCAGCCATGATCAATCCAGTGTTCCTCTATGAACAAT 1453
1355 GTCATGATTTTCAATAAAGCGGCACTCTGTGATGATATAAGTCCGTATATACCCAGTTGCAGGA 1414
1454 CGCCCAATGATGATCAAAACGGGATGAAATTTATCAATTTTACAAATTTTACAGAGTGTGGATCAT 1513
1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGGTGGATCAT 1474
```


;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655

;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 26.5%; Score 718.6; DB 13; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAACAAATGTCGAAGGCTGAAATATCTCAAGAAATGTTGGAATCCAAAC 328
DB 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAACCTACAAAGACTTGTCTTTCAAT 274
QY 329 AATGTGATCATTCTCAATGGCTTGGCCAACTCCAGTTATCATACCTTCTCTTTGGAT 388
DB 275 AGCTGTATTCCCTTTTGGGTTTCATGAAGACTGGGATTTTCAAACCTCTTCTTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATATTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGCGAGGCTGCTCTTGGAGCCAAAGACCACATCTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAAGATTGTGGCCAGTATCTTACCAAGAGAGATGAA 505
DB 395 GACTTAAACAAAAATTTAAGAAGATTTATTGGCCCTGCTGCAAGGAACGGGTGAATTA 454
QY 506 TCGAAGTGGCTGGAAAGACATCTCTGAAGATGCTGCTTAATTTTCATCAAGTACTTAAG 565
DB 455 TGTAATATTAGCTGGGAAGAGTGCCAAATACAGAATGTGCAAAATTTTCATCAGAGTACTCAG 514
QY 566 GCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTTCATCCAATTTGCACC 625
DB 515 CCTATACAAAATCATATATGTTGTGGAAGTGGAGCATTTTCATCCATATGTGGG 574
QY 626 TACATTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCAATTT 685
DB 575 TATATTGATCTTGGAGTCTCAAGAGGATATTATTTCAAACCTAGACACACATAATTTG 634
QY 686 GAAAACGGCCCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTCATCTCTCAGCAGCCTTTTGTCTTCAATATGACA 694

;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3430RIC104
;; CURRENT APPLICATION NUMBER: US/10/176,758
;; CURRENT FILING DATE: 2002-06-21
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 347
;; LENGTH: 3871
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-176-758-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGAAGAACATGTGCAAGGCTGAAATATCTACAAAGAAATGTGGAATCCAAC 328
DB 215 ACTTTGAAGCAAAATATCCAGACTCAAGCTAAACCTACAAAGACTTGTCTCTTCAAT 274
QY 329 AATGTGATCACTTTCATGCTTGGCCACAGCTCCAGTTATCATACCTTCTTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGTTTCATCAGAGACTGGATTTCAAACTCTCTCTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTGGAACAAAGGATCATATATTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGGAGGCTGCTCTTGGGAGCCAAAGACCACATCTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAGGATTTTCAAAGATTGTGGCCAGTATCTACACAGAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAAGATTTATTTGGCTGCTGCAAGGAACGGGTGGAATTA 454
QY 506 TGAAGTGGCTGGAAAGACATCTCTGAAAGATGTCTAATTTTCATCAAGTACTTAAAG 565
DB 455 TGTAAATAGCTGGGAAGATGCCAATACAGATGTGCAATTTTCTACAGACTTTCAG 514
QY 566 GCATATAATCAGACTCACTGTGACGCTGTGAAAGGGGGCTTTTCATCCAAATTTGCACC 625
DB 515 CCTATAACAAAACCTCACATATATGTGTGGAACCTGGAGCATTTTCATCCAAATATGTGG 574
QY 626 TACATTGAATTTGGACATCATCTGAGCAATATTTTAAAGTGGAGACTCAATTTT 685
DB 575 TATATTGATCTTGGAGTCTCAAGGAGGATTTATATTTCAAACTAGACACACATAATTTG 634
QY 686 GAAACGGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTCGATCTCTCAGAGCTTTTGTCTCAGTAATGACA 694
QY 746 GATGGAATATATCTCTGGAACCTGAGCTGATTTTATGGGCGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGATACTGCAATCACT 754
QY 806 CGAATCTTGGGC-----ACCACCAACCAATCAGACAGAGCAGCATGATTTCCAGG 856
DB 755 CGATCCCTTGGGCCTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTGGCCACCTCATCTCAGAGAGTGACAAATCCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTATTTGAACTTTCTTCATACACAGACACCTCAATCCAGAT 874
QY 917 GATGCAAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGGAGAACTCTCGGAAAA 976
DB 875 GATGATAAAATATATTTCTTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA 934
QY 977 GCTACTCAGCTAGATAGTGCAGATATGCAAGATGATTTTGGAGGCGACAGAGTCTG 1036
DB 935 ACCATCTTTCTCGAGTTGGAGAGATTTGTAAGATGATGTAGGAGGCAACGCGAGCTG 994
QY 1037 GTCAATAAATGACAACTTCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGCTCCAAT 1096
DB 995 ATAAACAAGTGGACGACTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGAAGTGTAT 1054
QY 1097 GCATTGACACTCACTTTTGTGATGAACTGCAGGATGATTTCCCTTAATGAACTTTAAAGATCCT 1156

DB 1055 GGGGCAGATACCTTACTTTGATGAGCTTCAAGATATTTTATTACTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGGAGTGTTTACGACTTCCAGTAAATTTTCAAGGATCAGCC 1216
DB 1115 AGAAATCCTGTAGTATATGGAGTCTTTTACTCAACCTCAGCTCCATCTTCAAGGCTTGTCT 1174
QY 1217 GTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATATCCACACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCATTTTAAATGGTCCATATGCTCATAAG 1234
QY 1277 GATGAACCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATATCCACGGCAGGA 1336
DB 1235 GAAAGTCAGACCACTCGTTGGTGCAGTATGATGGGAGAAATTCCTTATCCACGGCTGTCT 1294
QY 1337 ACTTGTCCACGAAACATTTTG--GTGGTTTGTACTCTACAAAGGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCCAGCAAAACCTATGACCCACTGATTAACTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAAGTCATCCAGCCATGTACAATCCAGTGTTCCTTATGAACAAT 1453
DB 1355 GTCATCAGTTTCATAAAGCGGCATCTCTGTGATGTATAGTCCGTATATCCAGTTGCAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACCGGATGTAATTTTCAATTTTACAAATTTTGTCTGAGACCGA 1513
DB 1415 GGCACCACTTTCAAGAGAAATCAATGTGATTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATCGAGAAGATGGACAGTATGATTTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTCATTGCAAGAGATGGCCAGTACGATGATGATTTTCTTGGAAACAGACATTTGGAATCTGTC 1534
QY 1574 CTTAAAGTAGTTTCAAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCTCTGCGAA 1633
DB 1535 CTCAAAGTTGTGAGCATTTCAAAGGAAAGTGG--AATATGGAAGAGTGTCTGCTGGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGTCTATTTTACGCAATGTGAGCTTTTCCACTAAGCAG 1693
DB 1592 GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTCTGAAGCAG 1651
QY 1694 CAACAACTATATATTTGTTTCAACGCTGGGTTGCCAGCTCCCTTTTACACCGGTGCTAT 1753
DB 1652 CAACAAATTTGATATTTGTTTCCGAGATGATTTGTTGAGTCTCTCTGCAAGATGGCAC 1711
QY 1754 ATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCCCGCAGACCCCTTACTGTCTGGGAT 1813
DB 1712 ACTTATGGGAAGCTTGGCAGACTGTTGTCTTGCAGAGAGACCCCTACTGTGCTGGGAT 1771
QY 1814 GGTTCGTGATGTTCTCGCTATTTTCCCACTGCAAGAGAGACGCAAGACGACAGATATA 1873
DB 1772 GGAATGCTGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCAAGATGTA 1831
QY 1874 AGAATGGAGACCCCACTGACTGCTCTGAGCTTTACAGCTTACACCATGATATACCATGGCCAC 1933
DB 1832 AATATGGGACCCCACTGACTGCTGAGCATGCAAGACAGCATTTAGTCTATG--AA 1889
QY 1934 AGCCTCGAAGAGAGAAATCATCTATGTTGAGAAATAGTAGCACATTTTGGAAATGCACT 1993
DB 1889 ACTGCTGATGAAAGTGTATTTTGGCATTTGAATTTTAACTCAACCTTTCTGAAATGATA 1948
QY 1994 CCGAAGTCGCAAGAGAGGCTGCTTATTTGGCAATTTCCAGAGGCGGAATGAAAGGCGGAAA 2053
DB 1949 CCTAAATCCCAACAAAGCAACTATTTAAATGGTATATCCAGAGGCTCAGGCGGATGAGCATCGA 2008
QY 2054 GAACAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTCTCTAGTCTAGT 2113
DB 2009 GAGGAGTTGAAGCCCGGATGAAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAAT 2069
QY 2114 CTACAAACAGAGGATTCAGGCAATTTACTCTGCAATGCGGTGGGAACATGGGTTTCATACAA 2173
DB 2069 TTGCAGAAAGAGGATTTCTGGAGTGTATTTACTGCAAAAGCCAGGAGACACTTTTCATCCAC 2128
QY 2174 ACTCTTCTTAAGGTAAACCTCGGAAATCATTTGACACAGAGCATTTTGGAAAGAACTTCTTCAT 2233

Db 2129 ACCATAGTGAAGCTGACTTTGATGTCAATGTAGATGAACAGATGGAATAATACCCAGAGG 2188
QY 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 8

US-10-175-737-347
; Sequence 347, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-737-347

Query Match 26.5%; Score 710.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGGAAGAACAAATGTGCCAAGGCTGGAATTTATCTCAAGAGAAATGTGGAATCCAAC 328
Db 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 274
QY 329 AATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGCTTATCATACACTTCTCTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGGATTTCAAACTCTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATATTTCAATTCGACCTGGTT 448
Db 335 GAGGAAGAGGAGGCTGTCTTGGGAGCCAAAGACCAATCTTCTACTCACTCTGGTT 394
QY 449 AA--TATCAAGGATTTTCAAGAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTAAGAGATTTATTTGGCTGTCTGCAAGGAAGCGGTTGAATTA 454
QY 506 TGCAAGTGGCTGGGAAGAGACATCTGAAAGATGTGCTAAATTTTCATCAAGTACTTAAG 565
Db 455 TGTAAATAGCTGGGAAGATCCCAATACAGATGTGCAAAATTTTCATCAGTACTTCAG 514
QY 566 GCATATAATCAGACTCATTGTCGCTGTGGAACGGGGCTTTTTCATCCAATTTGCACC 625
Db 515 CCTATAACAAAATCACATATATGTGTGGAACCTGGAGCATTTTCATCCAATATGTGG 574
QY 626 TACATTGAAATTTGGACATCATCTCAGGACATATTTTAACTGGAGACTCAATTTT 685
Db 575 TATATTGATCTTGGAGCTTCAAGGAGGATATTTATTTCAAACTAGACACACATATTTG 634
QY 686 GAAACGGGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGTCTCAGAGCATCCCTTTAATA 745
Db 635 GAGTCTGGCAGACTGAATGTCTTTCGATCTCTCAGAGGCTTTTGTCTCAGTATGACA 694
QY 746 GATGAGAAATTAATCTCTGGAACATGAGCTGATTTTATGGGGGAGACTTTTGTCTATCTTC 805

Db 695 GATGAGTACCTCTACTCTGGAAACAGCTTCTGATTTTCTTGGCAAGATACATGCTCACT 754
QY 806 CGAACTCTTTGGGC-----ACCACCAACCCCAATCAGGACAGAGCAGCATGATTTCCAGG 856
Db 755 CGATCCCTTTGGCCTACTCATGACCAACCACTATCATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTTCATTAGTGCCCACTCATCTCAGAGAGTGAACATCTGAA 916
Db 815 TGGCTCAATGGAGCAAAATTTATTGGAATTTCTTTCATACACAGACACCTTACATCCAGAT 874
QY 917 GATGACAAAGTATACCTTTTCTTCCGTGAAATGCAATAGATGGAGAGACATCTCGAAA 976
Db 875 GATGATAAAATATATTTCTTCTTCTGTAATCATCTCAAGAGAGGAGTACCTCCGATAAA 934
QY 977 GCTACTCAGCTAGATAGTAGTGCAGATATGCAAGATAGCTTTTGGAGGGGACAGAAAGTCTG 1036
Db 935 ACCATCCCTTCTCGAGTTTGGAGAGATTGTAGAAATGATGTAGGAGGACACAGCGAGCTG 994
QY 1037 GTGAATAAATGGACAAACATTTCTCAAAGCTCGTCTGTATTTGCTCAGTGCAGGTCCAAT 1096
Db 995 ATAAACAAGTGGAGGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTGGAAGTAT 1054
QY 1097 GGCATTGACACTCAATTTTGTAGAACTGCAGAGATGATTTCTTAATGAACTTTTAAAGATCT 1156
Db 1055 GGGGAGAGATCTTACTTTGATGAGCTTCAAGATATTTTACTCTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATATGAGTCTTTTACTCAACAGCTCCATCTTCAAGGGCTCTGCT 1174
QY 1217 GTGTGTATGTATAGCATGATGTGAGAAAGGGTGTCTTGTGTCTCATATGCCACACAGG 1276
Db 1175 GTTGTGTATAGCATGCTGACATCAGAGCAGTTTAAATGGTCCATATGCTCATAG 1234
QY 1277 GATGAACCCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCCTATTCACGGGCAGGA 1336
Db 1235 GAAAGTGAGAGACCATCTGTTGGGTGCAATGATGGGAGAAATTTCTTATTCACGGGCTGCT 1294
QY 1337 ACTTGTCCCAAGCAAAACATTTG---GTGGTTTGTGCTCTACAAAGGACCTTCTCTGATCAT 1393
Db 1295 ACATGTCAAGCAAAACCTATGACCCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGAACAAT 1453
Db 1355 GTCATCAGTTTCAATAAGCGGCACCTCTGTGATGTATAGTCCGTATATCCAGTTCAGGA 1414
QY 1454 CGCCCAATAGTGTATCAAAACCGGATGTAATTTATCAATTTTACAAATTTGCTGTAGACCGA 1513
Db 1415 GGAACCAAGCTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGCAGAAAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCTTT 1573
Db 1475 GTCATGTCAGAAAGATGGCCAGTACGATGTAATGTTTCTTGAACACAGATTTGGAACCTGTC 1534
QY 1574 CTAAAGTAGTCTCAATTTCTTAAGAGACTTGGTATGATTTTAGAAGAGGTTCTGCTGAA 1633
Db 1535 CTAAAGTTGTGACATTTCAAGGAAAGTGG---AATATGGAAGAGTAGTGTCTGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAAACCGACTGCTATTTTACGAATGGAGCTTTCACATTAAGCAG 1693
Db 1592 GAGTTGACAGATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTGCTCTGGAAGCAG 1651
QY 1694 CAAACAATATATATTTGTTCAACGGCTGGGTTTCCCGAGCTCCCTTTTACCGGGTGTAT 1753
Db 1652 CAACAATTTGATGTTGTTCCCGAGATGATTTAGTTCAGTCTCTCTGACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGAGAGCCCTTACTGTGCTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTCCAGAGACCCCTTACTGTGCTGGAT 1771
QY 1814 GGTTCGATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAACAGACGACATATATA 1873

Db 1772 GGAATGTCATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGGAGACCCACTGCTCACTCTTCAGACTTACACCATGATAATCACCATGCCAC 1933
Db 1832 AATATGGGACCCATCACCCTGCTGGACATCGAAGACAGCATTTAGTCATG---AA 1888
Qy 1934 AGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGCACATTTTTTGGAAATGCAGT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTTGGCAATTTAACTCAACCTTTCTGGAAATGTATA 1948
Qy 1994 CCGAAGTCGAGAGACCGCTGCTCTATTTGGCAATTTCCAGAGGGAATGAAGAGCGAATA 2053
Db 1949 CCTAAATCCCAACAGCACTATTAAATGGTATATCCAGAGGTCAGGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGAGTGATGATCATATCATCAGACAGATCAAGGCCCTCTGCTAGCTAGT 2113
Db 2009 GAGGAGTTGAGCCCGATGAAGAAATCATCAAAACGGAATATGGGCTACTGATTTGAAGT 2068
Qy 2114 CTACAAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTCATAAA 2173
Db 2069 TTGCAGAAGAGATTTCTGGATGTATTACTGCAAGCCCGAGGACACACTTTTCATCCAC 2128
Qy 2174 ACTCTTCTTAAGTAACCTCGAAGTCATTTGACACAGAGCATTTGGAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGAATGTCTTGAATGTCTTGAATGAACAGATGGAAATACCCAGAGG 2188
Qy 2234 AAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 9

US-10-173-706-347
; Sequence 347, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
Qy 269 AATGGAAGAACATGTGCGAAGGCTGAAATTTATCTTCAAGAAATGTTGGAATCCAAC 328
Db 215 ACTTTGAAGCAAAATATTTCAAGACTCAAGCTAACTCAAGAACTTGTGCTGCTTTCAAT 274
Qy 329 AATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGCTTATACACTTCTCTTTGGAT 388
Db 275 AGCTGTATTTCCCTTTTGGGTTTCATCAGAGGAGTGGATTTTCAAACTTCTCTTAGAT 334
Qy 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGAGATCACATATTTTTCATTCGACCTGGTT 448

Db 335 GAGGAAGAGGCGAGCTCTCTTGGAGCCAAAGACCACATCTTTCTACTCAGTCTGGTT 394
Qy 449 AA---TATCAGGATTTTCAAAAGANTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCCCTGCTGCAAAAGGAACGGGTGGAATTA 454
Qy 506 TGCAAGTGGGCTGGAAGAGACATCTGAAAGAAATGTGCTAAATTTTCAAGAGTACTTAAAG 565
Db 455 TGTAAATTTAGCTGGGAAGATGCCAATACAGAAATGTGCAAAATTTTCAAGAGTACTTCAAG 514
Qy 566 GCATATAATCAGACTCACTTGTAGCCCTGTGGAAAGGGGGCTTTTTCATCCAAATTTGCACC 625
Db 515 CCTATAACAAAACCTCACATATATGTGTGGAACTGGAGCATTTTATCCCAATATGTGGG 574
Qy 626 TACATTTGAAATGGACATCATCTGAGGACAAATATTTTAAAGCTGGGAACTCACATTTT 685
Db 575 TATATTGATCTTGGAGTCTCAAGAGAGATATTTATTTCAAACTAGACACACATATTTT 634
Qy 686 GAAAACGGCCCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
Db 635 GAGTCTGGCAGACTGAAATGTCCTTCGATCCTCAGCAGCCTTTTGTCTCAGTAATGACA 694
Qy 746 GATGAGAAATTTATCTCTGGAACTGCGAGCTGATTTTATGGGGCGAGACTTTTGTCTATCTTC 805
Db 695 GATGAGTACCTCTACTCTGGAAACAGCTTCTGATTTCTTGGCAAGAGATCTGCAATTCAT 754
Qy 806 CGAAGTCTTGGGCTCTCTGAGGAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTGAGG 856
Db 755 CGATCCCTTGGGCTCTCTGAGGAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTGAGG 814
Qy 857 TGGCTCAATGATCAAAAGTTCATTTAGTGGCCCACTCTCATCTCAGAGAGTGAACAATCCTGAA 916
Db 815 TGGCTCAATGAGCAGCAAAATTTTGTGAACTTCTTTCATACAGACACCTCAATCCAGAT 874
Qy 917 GATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAAACACTCTGGAATA 976
Db 875 GATGATAAATATATTTCTTCTTCTCGTGAATCATCTCAAGAAAGGAGTACCTCCGATAAA 934
Qy 977 GCTACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTTGGAGGCGACAGAGTCTG 1036
Db 935 ACCATCTCTTCTCGAGTTGGAAAGTGTGTAAGAAATGATGTAGGAGGACACGCGACCTG 994
Qy 1037 GTGAATAAATGGACAAACATTCCTCAAGCTCTGCTGATTTGCTCAGTCCAGGTCCTCAAT 1096
Db 995 ATAAACAGTGGACGACTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGGAGTGT 1054
Qy 1097 GGCATTGACACTCAATTTTGTATGAACTGAGATGTATTTCTTAATGAACCTTTTAAAGATCT 1156
Db 1055 GGGGCGAGATCTTACTTTGATGAGCTTCAAGATATTTATTTTACTCCCCACAGAGATGAA 1114
Qy 1157 AAAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACTTTTCAAGGAGTACGCC 1216
Db 1115 AGAAATCCTGTAGTATATGAGTCTTTACTTCAACAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATGTATAGCATGAGTGTGTGAGAGGTTTCTTGTGCTCATATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGGCTGACATCAGACAGTTTTATATGTTTCCATATGCTCATAG 1234
Qy 1277 GATGAGCCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCTTATCCAGCGGCGAGA 1336
Db 1235 GAAAGTCAGACACCATCGTTGGGTGAGTATGATGGAGAAATTCCTTATCCAGCGCTGCT 1294
Qy 1337 ACTTGTCCAGCAAAACATTTG---GTGGTTTTCAGTCTTCAAAAGGAGCTTCTCTGATCAT 1393
Db 1295 ACATGTCAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTGCAAGAAAGTCCATCCAGCCATGTACAAATCCAGTGTGTTTCTTATGAACAAT 1453
Db 1355 GTCATCAGTTTTCATAGCGGCACTCTGTGATGTATAGTCCGTATATCCAGTTGACAGA 1414
Qy 1454 CGCCCAATAGTGTATCAAAACGGAGTGAATTTATCAATTTTACAAATTTGTGTAGACCGA 1513

Db 1055 GGGGAGATACCTTACTTTGATGAGCTTCAAGATATATTTTACTCCCAAGAGATGAA 1114
Qy 1157 AAAAATCCAGTGTCTATATGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCGTGTATGAGTGTCTTACTACACAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGGTCCATATGCCCCACAGG 1276
Db 1175 GTTGTGTGTATAGCATGCTGACATCAGAGCAGTGTATTAAGTCCATATGCTCATAG 1234
Qy 1277 GATGGACCAACATATCAATGGGTGCTTATCAAGGAAGTCCCTATCCAGCGCCAGGA 1336
Db 1235 GAAAGTGCAGACCATCGTTGGGTGAGTATGATGGGAGAAATCTTTATCCAGGCGTGT 1294
Qy 1337 ACTGTGCCAGCAAAACATTTG---GTGGTGTGACTCTACAAGAGACCTTCTCATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTTATGACCACCTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTGCCAAGAGTCAATCCAGCATGTCAATCCAGTGTCTTCTATGAACAAT 1453
Db 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATGTATAGTCCGTATACCCAGTTGCAGA 1414
Qy 1454 CGGCCAATAGTGAATCAAAACGGATGAATTAATTAATTAACAATTTGCTGTAGACGA 1513
Db 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTACAGACTGACACAGATAGTGTGGATCAT 1474
Qy 1514 GTGGATGCAGAGATGCAGATGATGTATGTTTATCGGAACAGATGTTGGGACCGTT 1573
Db 1475 GTCAITGCAGAGATGGCCAGTACGATGAATGTTTCTTGGAAACAGCAATGGAACTGTC 1534
Qy 1574 CTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGTCTCTGCTGAA 1633
Db 1535 CTCARAAGTTGTCAGCATTTCAAAGGAAAGTGG---AATATGGAAGAGGTAGTGTGGAG 1591
Qy 1634 GAAATGACAGTGTTCGGGAAACGACTGCTATTTACGAATGAGAGCTTCCACTAAGCAG 1693
Db 1592 GAGTTGCAGATATTCAGGACACTCATCAATCATCTTGAACATGGAATGTCTCTGAAGCAG 1651
Qy 1694 CAACAACTATATATGTTTCAACGGCTGGGTGCCCCAGCTCCCTTTACACCGGTGTAT 1753
Db 1652 CAACAACTGTACATGTTCCGAGATGATTAAGTTTACGCTCTCTTGGACAGATGGCAG 1711
Qy 1754 ATTTACGGGAAAGCGTGTCTGAGTGTGCTCGCCGAGACCCCTTACTGTCTTGGGAT 1813
Db 1712 ACTATGGAAGAGCTTGGCAGAGCTGTGTCTTGGCAGAGACCCCTACTGTGCTGGAT 1771
Qy 1814 GGTTCGCAATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGACGACGAAGATATA 1873
Db 1772 GGAATGCAATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGA 1831
Qy 1874 AGAATGAGAGACCCACTGACTCACTGTTCCAGACTTACACCATGATAATCACCATGCCAC 1933
Db 1832 AAATATGGGACCCCAATCACCAGTCTGGGACATCGAAGACAGCATTAGTATG---AA 1888
Qy 1934 AGCCCTGGAAGAGAGATCATATATGTTGTAGAGATAGTAGCACATTTTTTGAATGCACT 1993
Db 1889 ACTGCTGATGAAAGAGTATTTTGGCAATTGAATTAATCACTCAACCTTCTGGAATGATA 1948
Qy 1994 CCGAGTCCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGGAATGAAGAGCGAATA 2053
Db 1949 CCTAAATCCCAAGCAACTATTAATGGTATATATCCAGAGGTGAGGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGT 2113
Db 2009 GAGGATGGAAGCCGATGAAGAAATCATCAAAACGAATATGGGCTACTGATTCGAAGT 2068
Qy 2114 CTACACAGAGAGATTCAGCAATTAACCTCTGCGCATGCGGTGGAAACATGGGTTTCACAA 2173
Db 2069 TTGCAAGAGAGATTCCTGGATGTATTACTGCAAAAGCCAGGACACACTTTTCATCCAC 2128
Qy 2174 ACTCTTCTTAAGTAACCCCTGGAAGTCAATGACAGAGCAATTTGGAAGAACTTCTCAT 2233
Db 2129 ACCATAGTGAAGCTGACTTTGAATGTGATTTGAATGAACAGATGGAATAATCCAGAGG 2188

Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 11

US-10-175-752-347
; Sequence 347, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-752-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

Qy 269 AATGGGAGACATGTGCCAAGCTGCAATTAATCTCTCAAGAAATCTCGAATCCCAAC 328
Db 215 ACTTGAAGCAAAATATTCGAAGTCAAGTAACTACAAAGACTTCTGCTTCAAT 274
Qy 329 AATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCTTTGGAT 388
Db 275 AGCTGATTCCTTTTGGGTTTCATCAGAGACTGGATTTTCAACTCTTCTTAGAT 334
Qy 389 GAGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCAATTCGACTGTT 448
Db 335 GAGGAAAGAGCAGGCTGCTCTGGGAGCCAAAGACACACATCTTTCTACTCAGTCTGTT 394
Qy 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTAAAGAAATTTATTTGGCTCTGCAAAAGGAACGGGTGGAATTA 454
Qy 506 TGCAAGTGGGCTGAAAGAGACATCTGAAAGATGTGTAATTTTCAAGGCTACTTAAG 565
Db 455 TGTAAATAGCTGGGAAGATGCCAATACAGAAATGTGCAATTTTCAACAGAGTACTTCA 514
Qy 566 GCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGGCTTTTCAATTCGACTGACC 625
Db 515 CCTATAACAAAACCTCATATATGTGTGGAACTGGAGCATTTTCAATTCGAATATGGG 574
Qy 626 TACATTGAATTTGACATCATCTTGGAGCAATATTTTAAAGCTGGGAACCTCACATTTT 685
Db 575 TATATTGATCTGTAGTCTACAAGGAGATATTTATTTCAAACTAGACACATATTTG 634
Qy 686 GAAACCGCCCTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTATA 745
Db 635 GAGTCTGCAGACTGAAATGTCTTTCGATCTCTCAGCAGCCCTTTTGTCTCAGTAATGACA 694
Qy 746 GATGAGAAATTAATCTCTGGAACCTGCAAGTGAATTTTATGGGGCGAGCTTTGCTATCTTC 805

Db 695 GATGAGTACTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATACTGCATCACT 754
Qy 806 CGAACTCTTGGGC-----ACGACCCCAATCAGACGAGCAGCAGATGATTCAGG 856
Db 755 CGATCCCTTTGGGCTACTCATGACCACTACTCATCAGAACTGCAATTTTCAGAGCACTAC 814
Qy 857 TGGCTCAATGATCCAAAGTTTCATAGTGCCCACTCATCTCAGAGAGTGCACAACTCTGAA 916
Db 815 TGGCTCAATGAGCAAAATTTATTTGAACTTTCTTATACCAAGACCTTCAATCCAGAT 874
Qy 917 GATGACAAAGTATPACTTTTCTTCCGTGAAATGCAATAGATGAGAACTCTGGAAAA 976
Db 875 GATGATAAAATATATTTCTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAA 934
Qy 977 GCTACTCAGCTAGAAATAGTCAGATATGCAAGATGACTTTGGAGGGCAGCAAGATCTG 1036
Db 935 ACCATCTTTCTCGAGTTTGAAGAGTTTGTAAAGATGATGTAGGAGGACAAGCGCAGCTG 994
Qy 1037 GTGAATAAATGGAACACATTTCTCAAGCTCTGATCTGATTTGCTCAGTGCAGGTCCAAAT 1096
Db 995 ATAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATCTCTGGAAGTGAT 1054
Qy 1097 GGCATTGACACTCATTTTGTGAATCTGAGGATGTATCTTAATGAATCTTAAAGATCCT 1156
Db 1055 GGGCAGATACTTCTTGTGAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
Qy 1157 AAAATCCAGTTGATATGAGGTTTACGACTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCCTGTAGTATGAGCTTTTACTCAACCCAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGATGATGATGATGATGATGAGAGGGTGTCTTGGTCCATATGCCACAGG 1276
Db 1175 GTTTGTGTATGAGCATGCTGACATCAGACGAGTTTAAATGGTCCATATGCTCATAG 1234
Qy 1277 GATGACCCCAATCATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGGCAGGA 1336
Db 1235 GAAAGTGACAGCATCTGTTGGGTGAGTATGATGGAGAAATCTTATCCAGCGCTGT 1294
Qy 1337 ACTGTGCCAGCAAAACATTTG---GTGGTTTTGACTTCAAAAGGACCTTCTCGATGAT 1393
Db 1295 ACATGTCCAAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATTAACCTTTGCAAGAGTATCCAGCCATGTCAATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATGTATAAGTCCGTATACCAGTTGCGGA 1414
Qy 1454 CGCCCAATGATGATCAAAAGCGATGTAAATTTATCAATTTACACAAATGCTGCTAGACCGA 1513
Db 1415 GGACCAAGTTCAAGAGATCAATGTGGATTCAGACTGACAGATAGTGGTGGATCAT 1474
Qy 1514 GTGGATGCAAGAGATGACAGATGATGATGTTTATCGGAACAGATGTTGGGACCGTT 1573
Db 1475 GTCAATGCAAGAGATGGCCAGTACGATGTAATGTTTCTTTGGAACAGACATTCGAACTGTC 1534
Qy 1574 CTTAAAGTATGTTCAATCTTAAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA 1633
Db 1535 CTCAAAGTTGTGAGCATTTTCAAGGAAAGAGTGG---AATATGGAAGAGGAGTGTCTGGAG 1591
Qy 1634 GAAATGACAGTTTCTTGGGAACCGACTGCTATTTTCCAGCAATGGAGTTTCCACTAAGCAG 1693
Db 1592 GAGTTGACAGATATTCAGACACTCATCAATCATCTTGAACATGGAATGTCTCTGAAGCAG 1651
Qy 1694 CAACAACTATATATTTGGTTTCAACGGCTGGGGTTGGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1652 CAACAACTGATCATTTGGTTCCGAGATGGATGATTTAGTTCAGTCTCTCTTGCACAGATGCGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTGCTGCGCCGAGACCTTTACTGTGCTGGAT 1813
Db 1712 ACTTATGGGAAGAGTTGCGACACTGTTGTCTTGGCAGAGACCTCTACTGTGCTGGAT 1771
Qy 1814 GGTCTGCAATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGACGCAAGATATA 1873
Db 1772 GGAATGCAATGCTCTCGATATGCTCTCTACTTCTTAAAGGAGAGCTAGAGCCCAAGATGTA 1831

Qy 1874 AGAATGGAGACCCCACTGACTCACTGTTTCAAGCTTACACCATGATATACCATGGCCAC 1933
Db 1832 AAATATGGCGACCCCAATCACCAGTGTCTGGGACATCGAAGACAGCATTAGTATGTCG---AA 1888
Qy 1934 AGCCCTTGAAGAGAGAAATCATCTATCTGTGTAGAGAAATAGTAGACATATTTTGGAAATGCACT 1993
Db 1899 ACTGCTGATGAAAGGTGATTTTGGCAATGAAATTTAACTCAACCTTCTCGAATGTATA 1948
Qy 1994 CCGAAGTCCGAGAGAGCGCTGCTTATTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 2053
Db 1949 CCTAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGCTCAGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGAGTGGATGATCATATCATCAGGAAGATCAAGGCTTCTGCTAGCTAGT 2113
Db 2009 GAGGAGTTTGAAGCCCGATGAAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT 2068
Qy 2114 CTCAACAAGAGGATTTAGGCAATTTACCTCTGCCATGGGTGGAACATGGTTTCATACA 2173
Db 2059 TTGCAAGAAGAGGATTTCTGGGATGTATTACTGCAAGCCCGAGGAGCACATTTTCATCCAC 2128
Qy 2174 ACTCTTTTAAGGTAACCTCGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGTCACTTTGAAATGCTCAATGAGATGACAGATGGAATAATACCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 12

US-10-176-482-347
; Sequence 347, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-482-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

Qy 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAAC 328
Db 215 ACTTTGAAGAAATAATTTCCAAAGACTCAAGCTAACCTAACAAAGACTTCTGCTTTCAAT 274
Qy 329 AATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCTCTTTGGAT 388
Db 275 AGCTGATATCTTTTGGGGTTTCATCAGAGGACTGGATTTTCAAACTCTTCTCTTAGAT 334
Qy 389 GAGGACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCAATTCGACTGTT 448

Db 335 GAGGAAAGAGCGAGGCTGCTCTTGGGAGCAAGAACACATCTTTCTACTGCTGCTT 394
Qy 449 AA---TATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAATAATTTAAGAGATTATTTGGCCCTGCTGCAAGGAAACGGGTGGAATTA 454
Qy 506 TGAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCAATCAAGGTACTTAAAG 565
Db 455 TGTAAATAGCTGGGAAAGATGCCAATACAGAATGTGCAAAATTTTCAATCAGAGTACTTCAG 514
Qy 566 GCATATAATCAGACTCAGCTGTGAGCTGTGAGAAACGGGGGCTTTTCAATCAATTTTGCACC 625
Db 515 CCCTATAACAAACTCACATATATGTGTGTGGAACCTGGAGCAATTTTCAATCAATATGTGGG 574
Qy 626 TACATTAATTTGAGCATCATCTGAGGACAAATATTTTAAAGCTGGAGACTCACAATTTT 685
Db 575 TATATTTGATCTTGGAGCTTACAAGGAGGATATTTATTTCAAACTAGACACACATATTTG 634
Qy 686 GAAAGCGGCGTGGGAAGAGTCCATATGACCCCTTAAGCTGTGACAGCATCCCTTTTAAATA 745
Db 635 GAGTCTGGCAGACTGAATGTCTTTTCGATCTCAGCAGCCTTTTGTCTCAGTAATGACA 694
Qy 746 GATGAGAAATTAATCTCTGAACTGAGCTGATTTTATGGGGGAGACTTTTGTCTATCTTC 805
Db 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGATAGTCAATCACT 754
Qy 806 CGAATCTTTGGG- - - - - ACCACCAACCAATCAGGACAGAGCAGCATGTTCCAGG 856
Db 755 CGATCCCTTTGGGCTTACTCATGACCACTTACTATCAGAACTGACATTTTCAAGAGCACTAC 814
Qy 857 TGCTCAATGATCCAAAGTTTCAATAGTGCCCACTCATCTCAGAGAGTGACAAATCTCTGAA 916
Db 815 TGGCTCAATGGGCAAAATTTATTGGAACTTTCTTCAATACCAGACACTTCAATCCAGAT 874
Qy 917 GATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGAGGAAACACTCTGGAATA 976
Db 875 GATGATAAAATATATTTCTTCTTCTGTAATCATCTCAAGAAAGGAGTACCTCCGATAAA 934
Qy 977 GCTACTCAGCTAGATAGTGCAGATATGCAGAAATGACTTTGGAGGACAGAGTCTG 1036
Db 935 ACCATCTTTCTGAGTGGAGAGTTTGTAAAGATGATGTAGGAGGAAACGCGAGCCTG 994
Qy 1037 GTGAATAAATGGACAACTTCTCAAGCTCTGATTTGCTCAGTGCAGGCTCCAAAT 1096
Db 995 ATAAACAAGTGGAGCACTTTCTTAAGGCCAGACTGATTTGCTCAATCTCTGGAAGTAT 1054
Qy 1097 GGCATTCAGACTCATTTTGAAGAACTGCGAGGATGATTTCTTAATGAACCTTTAAAGATCT 1156
Db 1055 GGGCGAGATACCTTACTTTGATGAGCTTCAAGATATTTTACTTCCCAACAAGAGATGAA 1114
Qy 1157 AAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATGAGGCTTTTACTACAACCAAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATGTATAGCATGATGTGAGAAAGGCTTCTTGGTCCATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGCTGCATCAGGAGTTTAAATGGTCCATATGCTATAAG 1234
Qy 1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAAAGAGTCCCCTATCCAGCGGACGAA 1336
Db 1235 GAAAGTGCAGACCACTGTTGGGTGAGTATGATGGGAGAAATCTTTATCCAGGGCTGTG 1294
Qy 1337 ACTTGTCCAGCAAAACATTTG- - - GTGGTTTGTACTTCAAGAGACCTTCTGTATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTTATGACCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTTCAAGAAAGTATCCAGGCCATGTACAATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATCAGTTTCAATAAGCGGCACTCTGTGATGTATAAGTCCGTATACCGAGTTCGAGA 1414
Qy 1454 CGCCCAATAGTATCAAAACGGAGTAAATTTATCAATTTACAAATTTGCTGTAGACGAA 1513
Db 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTACAGACTGACAGATAGTGTGGTATCAT 1474

Qy 1514 GTGATGCAAGAGATGGACAGTATGATCTATTTTATCGGAAACAGATGTTTGGGACCGTT 1573
Db 1475 GTCTTGCAGAGATGGCCAGTACGATGATTTTCTTGGAAACAGACATTTGGAATGTC 1534
Qy 1574 CTTAAAGTATTTCAATTTCTTAAGGAGACTTGGTATGATTTTAAAGAGAGTTCTGTGGA 1633
Db 1535 CTCAAAGTTGTGAGCATTTTCAAGGAAAGTGG- - - AATATGGAAGAGGTAGTGTCTGGAG 1591
Qy 1634 GAATGACAGCTTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTTCCACTAGCAG 1693
Db 1592 GAGTTGTCAGATATTTCAAGCACTCATCAATCATCTTTGAAACATGGAATTTGCTCTGAAGCAG 1651
Qy 1694 CAACAACATATATATTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1652 CAACAAATTTGATATTTGTTCCGAGATGATTTAGTTTCACTCTCTTGCACAGATGCGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTGCTGCCCGAGACCTTTACTGTGCTTGGAT 1813
Db 1712 ACTTATGGGAAGCTTTGGCAGACTGTTGTCTTGGCAGAGACCCCTACTGTGCTGGGAT 1771
Qy 1814 GGTTCGATGTTCTGCTATTTTCCCACTGCAAGAGACCGCAAGAGCAGCAAGATATA 1873
Db 1772 GGAATGCACTGCTCTCGATATGCTCTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGAGAGACCCACTGACTGCTTTCAGACTTTACACCATGATAATCACCATGGCCAC 1933
Db 1832 AAATATGGGACCCCATCACCAGTGTGGGACATCGAAGACAGCATTTAGTCATG- - - NA 1888
Qy 1934 AGCCTTGAAGAGAGATCATCTATGTTGAGAGATAGTAGACATTTTGGAAATGCACT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATTTAACTCAACCTTTCTGGAATGTATA 1948
Qy 1994 CGAAGTCCGAGAGCGCTGCTTATTTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 2053
Db 1949 CCTAAATCCCAACAGCAACTATTAAATGGTATATCCAGAGGTGAGGAGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGATGAGTATCATCATCAGGACAGATCAAGGCTTTCTGCTAGCTAGT 2113
Db 2009 GAGGATTTGAAGCCGATGAAAGATCATCAACGGATATTTGGCTACTGATTCGAGT 2068
Qy 2114 CTACAACAGAGAGATTCAGGCAATTTACCTTGCCTCAGGTTGGAAACATGGGTTTATACAA 2173
Db 2069 TTGCAGAAAGAGATTTCTGGGATGATTTACTGCAAAAGCCAGGAGCACACTTTTCATCCAC 2128
Qy 2174 ACTCTTCTTAAGTAAACCTGGAAGTCAATTCACACAGAGCATTTTGGAGAACTTTCTTCT 2233
Db 2129 ACCATAGTGAAGTCACTTTGAATGCTATTGAGAAATGAACAGATGGAATAATCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 13

US-10-176-757-347
; Sequence 347, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deanovs, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C86

```
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-347

Query Match          26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred No. 7,9e+205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAAACAATGTGCAAGGCTGAAATTTATCTCAAAAGAAATGTGGAATCCAAC 328
DB 1115 AGAAATCTGTAGTATATGGAGTCTTTACTACAAACGAGTCCATCTTCAAAGSCTCTGCT 1174
QY 1217 GTGTGTATGTATAGCATGATGTGAGAGAGGTGTCCTTGGTCCATATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTTAATGGTCCATATGCTCATAG 1234
QY 1277 GATGACCCAACTATCAATGGTGGCTTTATCAAGGAAGAGTCCCTATCCAGGCCAGGA 1336
DB 1235 GAAAGTCAGACCATCTGTTGGTGCAGTATGATGGGAGAAATTCCTTATCCAGCGCTGGT 1294
QY 1337 ACTTGTCCAGCAAAACAATTTG---GTGGTTTTGTACTCTACAAAGAGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCGAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATACTTTGCAAGAAAGTCATCCAGCCATGTCATCAATCAAGTGTTCCTTATGAACAT 1453
DB 1355 GTCATCAGTTTTATAAAGCGGCACCTCTGTGATGATTAAGTCGGTATACCCAGTTGACGA 1414
QY 1454 CGCCCAATAGTATGATCAAAACCGATGTAATTTATCAAAATTTACAAATTTGCTAGACGA 1513
DB 1415 GGCACCAAGTTCAAGAGAAATCAATGTGGATTAACAGCTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGAGAAAGATGGAAGTATGATTTATGTTTATCGAAACAGATGTTTGGGACCGTT 1573
DB 1475 GTCAATGAGAAAGATGGCCAGTACGATGTAATTTCTTGGAAACAGACATTTGGAATGTC 1534
QY 1574 CTTAAAGTAGTTTCAATTTCTTAAGAGACTTGGTATGATTTTAGAAGAGGTTCTGCTGAA 1633
DB 1535 CTCAAGTTGTGAGCATTTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTGAG 1591
QY 1634 GAAATGACAGTTTTTCGGGAACCGACTGCTATTTTCAAGCAATGGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTGACAGATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTGAAGCAG 1651
QY 1694 CAACAACTATATATTTGGTTTCAACGCGTGGGGTTCGCCAGCTCCCTTTACACGGGTGAT 1753
DB 1652 CAACAAATGTACATTTGTTCCCGAGATGGAATTTAGTTTCACTCTCTTCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGAGGTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTTGGGAT 1813
DB 1712 ACTTATGGAAGAGCTTGGCAGACTGTTGTCTTGCAGAGACCCCTACTGTGCTGCTGGAT 1771
QY 1814 GGTTCGATCTCTCGCTATTTTCCCACTGCAAGAGACGACCAAGAGACGACAGATATA 1873
DB 1772 GGAATGATGCTCTCGATATGCTCTTAAAGGAGAGCTTAGACGCCAAGATGTA 1831
QY 1874 AGAAATGAGAGACCACTGACTCAGCTGTTTCAAGCTTTACACCATGATTAATCACCATGGCCAC 1933
DB 1832 AAATATGCGACCAATCAACCCAGTGTGGGACATCGAAGACAGCATTTAGTCAATG---AA 1888
QY 1934 AGCCCTGAAGAGAGAAATCATCTATGTTGTTGAGAAATAGTAGACATTTTGGAAATGCAAT 1993
DB 1889 ACTGCTGATGAAAGGTGATTTTGGCATTTTAACTCAACCTTTCTGGAATGTATA 1948
QY 1994 CCGAGTGCAGAGAGCGCTGCTGCTATTTGGCAATTTCCAGAGCGGAAATGGAAGAGCAAAA 2053
DB 1949 CCTAAATCCCAACAGCAACTTAATTAATGGTATATCCAGAGGTGAGGAGTGAAGATCGA 2008
QY 2054 GAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTTCTGCTAGCTAGT 2113
DB 2009 GAGGATTTGAAGCCCGATGAAAGAAATCATCAAAAGGAAATATGGCTACTGATTCGAATG 2068
QY 2114 CTAACAAGAGAGATTCAGGCAATTAACCTGTCCATGCGGTGGAAACATGGGGTTCATACA 2173
DB 2069 TTGAGAGAGAGGATTTCTGGGATGTTTCTGCAAAAGCCCAAGGAGACACATTTTATCCAC 2128
QY 2174 ACTCTTTAAGGTAAACCTGGAAAGTCAATTCACACAGAGCAATTTGGAGAACTTCTTCAT 2233
DB 2129 ACCATAGTGAAGCTGACTTTGAATGTCTATTGAGAAATGAACAGATGGAATAATACCCAGAG 2188
```

QY 2234 AAAGATGATGATG 2246
DB 2189 GCAGAGCATGAGG 2201

RESULT 14

US-10-176-913-347
; Sequence 347, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P34301C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-913-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGAAGAACAAATGTGCCAAGGCTGAATATCTCCACAAAGAAATGTTGGAATCCAAAC 328
DB 215 ACTTGAAGCAAAATATCCAAAGCTCAAGCTTAACCTACAAAGACTTGTGCTGCTTCAAT 274
QY 329 AATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGTTATCAVACCTTCCCTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACCTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTGGAGCAAGAGATCAATATTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGGCGAGGCTGCTCTTGGGAGCCAAAGACCACATCTTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGATTTTCAAAGATTTGTGCCAGTATCTTACACCAAGAGATGAA 505
DB 395 GACTTAAACAAATTTTAAAGATTTATTTGGCTGTCTGCAAGAGGAGCGGTGGAATTA 454
QY 506 TGCAGTGGGCTGGAAAGACATCTCTGAAAGATTTGTCTAAATTTTCATCAAGGTACTTAAG 565
DB 455 TGTAAATAGCTGGGAAGATGCCAATACAGATGTGCAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCATTGTGACGCTGTGGAGCGGGGCTTTTCATCCTCAATTTGACCC 625
DB 515 CCTATAACAAAACTCACATATATGTGTGGAATCTGAGAGCATTTTCATCCAAATATGTGG 574
QY 626 TACATTGAATTTGGACATCATCTGAGGACATATTTTAAAGCTGGAGACTCACAATTTT 685
DB 575 TATATTGATCTTGGAGCTTCAAGGAGGATATTTATTTCAAACTAGACACATATTTG 634
QY 686 GAAACCGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTTCGATCTCTCAGCAGCTTTTGTCTTCAATATGACA 694
QY 746 GATGGAGATTTACTCTGGAATCTGAGCTGATTTTATGGGGCGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACTCTTCTCTGGAACAGCTTCTGATTTTCTTGGCAAGAGATGACTGCAATTCAT 754

QY 806 CGAACTCTTGGGC-----ACCACCAACCAATCAGGACAGAGCAGCATGATTCACGG 856
DB 755 CGATCCCTTGGGCTTACTCATGACCAACCACTACATCAGAACTGACATTTTCAGAGCATAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAACTCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTTATGGAACCTTTCTTATACACAGACACCTTACAATCCAGAT 874
QY 917 GATGCAAAAGTATCTTTTCTTCGTTGAAAATGCAATAGATGGAGAACACTCTGGAAA 976
DB 875 GATGATAAAATATATTTCTTCTTCTCGTGAATCATCTCAAGAGAGGAGTACCTCCGATAAA 934
QY 977 GCTACTCAGCTAGAAATAGGTAGATGCAAGATGACTTTGGAGGGGACAGAAAGTCTG 1036
DB 935 ACCATCCTTTCTCGAGTTGGAAGAGTTTGTGAAGATGATGTAGGAGGACACAGAGCTG 994
QY 1037 GTGAATAAATGACAAACATCTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAAT 1096
DB 995 ATAAACAAGTGGAGCAGCTTTTCTTAAGCCAGACTGATTTGCTCAATTTCTGGAAGTGT 1054
QY 1097 GGCATTGACACTCATTTTGTGAACTGAGGATGTATTCTTAATGAACCTTTTAAAGATCCT 1156
DB 1055 GGGCAGATATCTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGAGTACGCC 1216
DB 1115 AGAATCTCTGATGATATGAGAGTCTTTACTACACAGACTCCTCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGTATAGCATGAGTGTGTGAGAGGGTGTCTTGTGTCATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTTAATGGTCCATATGCTCATAAG 1234
QY 1277 GATGACCCAACTATCAATGGTGGCTTATCAAGAGAGAGTCCCTATCCAGGCGCAGA 1336
DB 1235 GAAAGTGAGAGCCTCGTTGGGTGAGTATGATGGAGAAATTTCTTATCCAGCGGCTGGT 1294
QY 1337 ACTTGTCCCAAGCAAAACATTTTGTGTTTGTACTCTACAAAGAGACTTCTCTGATGAT 1393
DB 1295 ACATGTCCAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTATCCAGCATGATCAATCCAGTGTTCCTTATGAACAAAT 1453
DB 1355 GTCATCAGTTTCAAAAGCGGCACTCTGTGATGTATAGTCCGTATATCCAGTTTCAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGGATGTAAATTTCAATTTTACAAATTTGCTGTAGACCA 1513
DB 1415 GGACCAACGTTCAAGAGAAATCAATGTGATTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGCAGAAAGATGAGCAGTATGATTTATGTTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTCATTTGCAAGAGATGGCCAGTACGATGTATGTTTCTTGGAAACAGACATTTGGAATGTC 1534
QY 1574 CTTAAAGTAGTTTCAAATTCCTAAGGAGACTTTGGTATGATTTTAGAAGAGGTTCCTGCGAA 1633
DB 1535 CTCAAAGTTGTGAGCATTTTCAAGAGGAAAGTGG---AATATGGAAGAGGTAGTGTGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGCTATTTTACGCAATGAGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTGACAGATTTTCAAGCAGCTCATCAATCATCTTGAACATGGAATTTCTCTCTGAAGCAG 1651
QY 1694 CAACCACTATATATTTTGGTTTCAACGGCTGGGGTGTGCCAGCTCCCTTTTACACGGTGTGAT 1753
DB 1652 CAACAAATGTACATTTGGTTCCGAGATGGATTTAGTTTCTGCTCTCTTCCACAGATGCCAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGAGAGACCTTTACTGTCTTGGGAT 1813
DB 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTGGCAGAGACCCCTACTGTGCTGCTGGAT 1771
QY 1814 GGTCTGCAATGCTCTCGCTATTTTCCCACTGCAAGAGACCGCACAGACACAGATATA 1873
DB 1772 GGAATGCACTCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831

QY 1874 AGAATGGAGACCCAGTCACTGTTAGAGTATACCATGATGATGAGTGGCCAC 1933
Db 1832 AAATATGGCGACCCAAATCAACCCAGTCTGGGATCGAAGACAGCATGATGATCATG---AA 1888
QY 1934 AGCCCTGAAGAGAGAGATCATCTATGCTGTAGAGATAGTAGACATTTTGGAAATGCACT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATGATTTAACTCAACCTTCTGGANGTATA 1948
QY 1994 CGAAGTCCGAGAGAGCGTGTGTTATGCGCAATTCAGAGGCGAAATGAAGAGCGAAAA 2053
Db 1949 CCTAAATCCCAACAGCAACTATTAAATGGTATATCCAGAGGTGAGGGATGAGCATCGA 2008
QY 2054 GAAGATCAGAGTGGATGATCATATCATCAGAGCATGAGGCCCTTCTGCTACGTAGT 2113
Db 2009 GAGGAGTTGAAGCCCGGAAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT 2068
QY 2114 CTACACAGAGGATTCAGGCAATTAACCTCTGCGATGCGGTGGAACATGGGTTTATACAA 2173
Db 2069 TTGCAAGAGAGATTTCTGGATGATTAATCTGCAAGCCAGGAGCACATTTTATCCAC 2128
QY 2174 ACTTCTTTAAGTAAACCTGGAAGTCAATTGACACAGAGCATTTTGAAGAATCTTTCAT 2233
Db 2129 ACCATAGTGAAGTCACTTTGAATGTCAATTGAGATGAACAGATGGAATAATCCAGAGG 2188
QY 2234 AAGATGATGATG 2246
Db 2189 GCAGAGCATGAG 2201

RESULT 15

US-10-180-552-347
; Sequence 347, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Destoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-552-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGAAGAACAAATGTGCAAGGCTGAAATTTATCTACAAAGAAATGTGGAATCCAAC 328
Db 215 ACTTTGAAGCAAAATATTCAGAGCTCAGCTTAACCTACAAAGACTTGTCTGCTTTCAAT 274
QY 329 AATGATCAGCTTCAATGGCTTGGCCAAACAGCTCAGTTATCAATCTTCTTTTGGAT 388
Db 275 AGCTGATTTCCCTTTTGGGTTTCATCAGAAGGACTGGAATTTTCAAACTCTTCTCTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATCTGGAGCAAGAGTACATATTTTTCATTCAGCTGTT 448
Db 335 GAGGAAGAGGAGGCTGTCTTTGGAGGCCAAAGACCATCTTTCTACTAGTCTGTT 394

QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCCTGCTGCAAGAGAACCGGTGGAATTA 454
QY 506 TCCAGTGGCTGGAAAGACATCTCTGAAGATGCTGCTAAATTTTTCATCAAGGTACTTAAG 565
Db 455 TGTAATTTAGCTGGGAAGAGTCCAAATACAGAAATGTCAAAATTTTTCATCAGATCTTCAG 514
QY 566 GCATATAATCAGACTCACTTGTACCGCTGTGAAACGGGGCTTTTTCATCCAATTTGCAACC 625
Db 515 CCTATAACAAACCTCACAATATATGCTGTGGAACCTGGAGCATTTTCAATCAATATGTTGG 574
QY 626 TACATTTGAAATTTGGACATCATCTCTGAGCAATATTTTAAAGCTGGAGAACTCAATTTT 685
Db 575 TATATTTGATCTTGGAGTCTCAAGGAGGATATTAATTTCAAACTAGACACACATAATTTG 634
QY 686 GAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTCACAGCATCCCTTTTAAATA 745
Db 635 GAGTCTGGCAGACTGAAATGTCTTTCATCTCAGCAGCCTTTTGTCTTCAGTAATGACA 694
QY 746 GATGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTGCTATCTTC 805
Db 695 GATGAGTACCTTCTCTGGAACAGCTTCTGATTTCTTGGCAAGATATCTGATCTACT 754
QY 806 CGAAGTCTTGGGC-----ACCACCAACCAATCAGACACAGACGATGATTTCCAGG 856
Db 755 CGATCCCTTGGCCTTACTCATGACCAACCTACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATTTAGTGCACCTCATCTCAGAGGTGACATCTGNA 916
Db 815 TGGCTCAATGAGCAAAATTTTATTTGGAACCTTTCTTATACCAAGACCTTACATCAAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTCGMAAA 976
Db 875 GATGATAAATATATTTCTTTCTGTAATCATCTCAAGAGGAGTACCTCCGATAAA 934
QY 977 GCTACTACGCTAGTAAGTCAAGATATGCAAGAAATGACTTTTGGAGGCGACAGAGTCTG 1036
Db 935 ACCATCTTTCTCGAGTTTGGAAAGTGTGTAAGAAATGATGAGGAGGACAAACGAGCCTG 994
QY 1037 GTGAATAATGGACAAACATTTCTCAAGCTGCTGATTTGCTCAGTGCAGCTCAAT 1096
Db 995 ATAAACAAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGAAGTGT 1054
QY 1097 GCGATTGACACTCATTTTGTGAACTGAGAGTGTATTTCTTAATCAACTTTTAAAGATCCT 1156
Db 1055 GGGGAGAGATCTTCTTTGATGAGCTTCAAGATATTTATTTACTCCCAAGAGATGNA 1114
QY 1157 AAAAAATCCAGTTGATATGAGTGTGTTACGACTTTTCAAGTAACTTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCTGTAGTATATGAGTCTTTTACTACAAACGAGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
Db 1175 GTTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1234
QY 1277 GATGACCCCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCTTATCCAGGCGAGGA 1336
Db 1235 GAAAGTGCAGACCATCTGTTGGGTGAGTATGATGGAGAAATTTCTTTATCCAGGCTGTG 1294
QY 1337 ACTTGTCTCCAGCAAAAACATTTG---GTGGTTTGTGACTCTACAAAGGAGCTTTCTCTGATGAT 1393
Db 1295 ACATGTCCAAGCAAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTCAATCCAGCATGATCAATTCAGTGTGTTCTTATGAAACAT 1453
Db 1355 GTCATCAGTTTCAATAAAGCGGCACTCTGTGATGATTAAGTCCGTATATCCAGTGTGAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGAGTAAATTTTCAATTTTACACAAATTTGCTGAGACCGA 1513
Db 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGGTGTGATCAT 1474

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search time 27.5 Seconds
(without alignments)
9265.116 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttatcatgatg.....aggctttttttcttcataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match' 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09774490/runat_03082003_095416_25322/app_query.fasta_1.2887
-DB=SwissProt_41 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1 1 33 @runat_03082003_095416_25322 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4201	85.8	771	1 SM3A_HUMAN	Q14563 homo sapien
2	4057.5	82.9	772	1 SM3A_RAT	Q63548 rattus norv
3	4035.5	82.4	772	1 SM3A_MOUSE	O08665 mus musculu
4	3784.5	77.3	772	1 SM3A_CHICK	Q90607 gallus gall
5	3207.5	65.5	778	1 S21B_BRARE	Q9W686 brachydanio
6	3075.5	62.8	860	1 S21A_BRARE	Q9W7J1 brachydanio
7	2247.5	45.9	777	1 SM3D_HUMAN	O95025 mus sapien
8	2245	45.9	761	1 SM3D_CHICK	Q90663 gallus gall
9	2125	43.4	764	1 SM22_BRARE	Q9W696 brachydanio
10	2118	43.3	749	1 SM3B_HUMAN	Q13214 homo sapien
11	2059.5	42.1	748	1 SM3B_MOUSE	Q62177 mus musculu
12	2044	41.7	775	1 SM3E_HUMAN	O15041 homo sapien
13	2020	41.3	775	1 SM3E_MOUSE	P70275 mus musculu
14	1993	40.7	785	1 SM3E_CHICK	O42337 gallus gall
15	1940	39.6	785	1 SM3F_HUMAN	Q13275 homo sapien
16	1928	39.4	785	1 SM3F_MOUSE	O88632 mus musculu
17	1894	38.7	751	1 SM3C_CHICK	O42236 gallus gall
18	1850.5	37.8	751	1 SM3C_MOUSE	Q62181 mus musculu

19	1849.5	37.8	751	1 SM3C_HUMAN	Q99985 homo sapien
20	1003	20.5	861	1 SM4D_MOUSE	O09126 mus musculu
21	981.5	20.0	862	1 SM4D_HUMAN	Q92854 homo sapien
22	966	19.7	834	1 SM4C_MOUSE	Q64151 mus musculu
23	959.5	19.6	837	1 SM4G_MOUSE	Q9W4H7 mus musculu
24	937	19.1	838	1 SM4G_HUMAN	Q9N9N9 homo sapien
25	889	18.2	832	1 SM4B_HUMAN	Q9NPR2 homo sapien
26	857.5	17.5	782	1 SM4B_MOUSE	Q62179 mus musculu
27	852.5	17.4	1030	1 SM6A_HUMAN	Q9H2E6 homo sapien
28	834.5	17.0	888	1 SM6A_MOUSE	O35464 mus musculu
29	831	17.0	766	1 SM27_BRARE	Q9VHX4 brachydanio
30	825.5	16.9	761	1 SM4A_HUMAN	Q9H3S1 homo sapien
31	802	16.4	760	1 SM4A_MOUSE	Q62178 mus musculu
32	792	16.2	771	1 SM1A_DROME	Q24322 drosophila
33	783.5	16.0	697	1 SM2A_SCHGR	Q9XZC8 schistocerc
34	780.5	15.9	706	1 SM2A_DROME	Q24323 drosophila
35	766.5	15.7	730	1 SM1A_SCHAM	Q26473 schistocerc
36	766	15.6	1093	1 SM5B_MOUSE	Q60519 mus musculu
37	760.5	15.5	930	1 SM6C_HUMAN	Q9H3T2 homo sapien
38	756.5	15.5	777	1 SM4F_MOUSE	Q92123 mus musculu
39	752.5	15.4	776	1 SM4F_RAT	Q92143 rattus norv
40	745.5	15.2	960	1 SM6C_RAT	Q9WTL3 rattus norv
41	742	15.2	931	1 SM6C_MOUSE	Q9WTL3 rattus norv
42	737.5	15.1	770	1 SM4F_HUMAN	O95754 homo sapien
43	737.5	15.1	1077	1 SM5A_MOUSE	Q62217 mus musculu
44	730	14.9	1074	1 SM5A_HUMAN	Q13591 homo sapien
45	723	14.8	887	1 SM6B_RAT	O70141 rattus norv

ALIGNMENTS

RESULT 1
SM3A_HUMAN
ID SM3A_HUMAN STANDARD; PRT; 771 AA.
AC Q14563;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
GN SEMA3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Matthes D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RT growth cone guidance molecules.";
RL Cell 75:1389-1399(1993).
RN [2]
RP SEQUENCE OF 1-37 FROM N.A.
RA Wessner J., Minx P., Hinds K., Strommatt C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-182 FROM N.A.
RA Rohlfing T., Tin-Wollam A.M., Duckels G.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
CC NEUROPILIN-1/PLEXIN-1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L26081; AAC65938.1; -;
 DR EMBL; AC004451; -; NOT_ANNOTATED_CDS.
 DR EMBL; AC004848; AAC78622.1; -;
 DR PIR; D49423; D49423.
 DR Genew; HGNC:10723; SEMA3A.
 DR MIM; 603961; -;
 DR GO; GO:0005576; C:extracellular; TAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; IG; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 771 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 580 664 IG-LIKE C2-TYPE.
 FT DOMAIN 727 769 ARG/LYS-RICH (BASIC).
 FT DISULFID 649 722 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 771 AA; 88889 MW; 9985F8D3EAD8456 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_HUMAN (1-771)

QY 200 ATGGGCTGTTAACTAGGATTCTGTCTTTCTGGGAGTATTACTTACAGCAAGAGCA 259
 DB 1 MetGlyTTPLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAGGCTGAAATATCTTACAAAGAAATGTTG 319
 DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCACAATGTGATCACTTTCAATGCTTGGCCACAGCTCCAGTTATCATCCCTTC 379
 DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGCAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTTC 439
 DB 61 LeuLeuAspGluLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTTAATATCAAGATTCTTCAAAAGATGTGTGGCCAGTATCTTACACCAAGAGA 499
 DB 81 AspLeuValAsnIleLysAspPheGlnIleValTTPProValSerTyrThrArgArg 100
 QY 500 GATGATGCACTGGCTGGAAAGACATCTCTGAAGATGTCTAAATTTTCATCAAGGTA 559
 DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATAATCAGACTCACTTGTCAGCTGTGGAAACGGGGCTTTTCATCCAATT 619

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGCAATATATTTTAACTGAGACTCA 679
 Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTGACACATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTGT 799
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAACCTCTTGGCACCACCACCATCAGGACAGACAGACATGATTCAGAGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCCTGAAGAT 919
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTTCTCCGTCAAAATGCAATAGATGGAGAACACTCTGGAAGAGCT 979
 Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGAAATAGGTGAGATATGCAAGATGACTTTTGGAGGCCACAGAGTCTGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACATCTCTCAAGCTCGTCTGATTTTCTCAGTGCCAGGTCCAAATGGC 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTGATGAACACTGCAGGATGATTTCTTAATCAACTTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCAGTTGTATATGAGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGTGTCCATATATGCCACAGGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAACATTTGGTGGTTTTCACCTCTACAAAGGACCTTCTCTGATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGCAACAAGTCACTCCAGCCATGACCAATCCAGTGTTCCTATCAACAATCCGCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 CGAGAGATGACAGATGATGTTTATGTTTATCGAACAGATGTTGGACCGTCTTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTAATTAGAAAGAGGTTCTCTGTGAAGAAATG 1639
 Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCATTAAGCAGCAACA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500

QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
 Db LeuTyIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GCGAAGCGTGTGCTGAGTGTTCCTCGCCGCGAGACCCCTTACTGCTTGGATGGTCT 1819
 Db GlyIysAlaCysAlaGluCysIleuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGCTATTTCCTCAACTGCAAGACGACAGACGCAAGATATAAGAAAT 1879
 Db AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgen 560
 QY 1880 GGAGACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
 Db GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTACACATTTTGGATGCTAGTCCGAG 1999
 Db GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGACGCTGCTGCTATTTCGCAATTCAGAGCGCAAAATCAAGAGCAAAAGAGAG 2059
 Db SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluArgLysGluGlu 620
 QY 2060 ATCAGAGTGATCATATCATATCATATCATATCATATCATATCATATCATATCATAT 2119
 Db IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGGATTCAGGAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2179
 Db GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGTGAACCTCGAAGTCAATTCACACAGAGCAATTTGGAAGAACTTCTTCATAAGAT 2239
 Db LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
 QY 2240 GATGATGAGATGGCTTAAGACCAAGAAATGTCCAATAGATGATGATGATGATGATGAT 2299
 Db AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGTCAGAGACTTATGTCAGCTCATACACACACACACACACACACACACACACAC 2359
 Db ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
 QY 2360 TTCTGTGAACAGTTTGGAAAAGGACCGCAAAACACCTCGGCAAGCGCAGGACATACC 2419
 Db PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
 QY 2420 CCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAACAGGAGG 2479
 Db ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
 QY 2480 ACCCAGCAATTTGAGAGCGCCACCGAGGAGTGC 2512
 Db ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

SM3A_RAT
 ID SM3A_RAT STANDARD; PRT; 772 AA.
 AC Q63548;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;

RX MEDLINE=97073089; PubMed=8915837;
 RT Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaagen J.;
 RA "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
 relationship to developing nerve tracts during neuroembryogenesis.";
 RL J. Comp. Neurol. 375:378-392(1996).
 CC -!- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
 CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
 CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
 CC NEUROFILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
 CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELENCEPHALIC
 CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
 CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
 CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
 CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
 CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
 CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
 CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
 CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
 CC SPINAL MOTONEURONS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL; X95286; CAA64607.1; -;
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG_LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 20 POTENTIAL.
 CC CHAIN 21 772 SEMAPHORIN 3A.
 CC DOMAIN 240 538 SEMA.
 CC DOMAIN 577 665 IG-LIKE C2-TYPE.
 CC DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 CC DISULFID 650 723 BY SIMILARITY.
 CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 772 AA; 8880 MW; 240907812PFFSF2D2 CRC64;

Alignment Scores:

Pred. No.: 9.28e-316 Length: 772
 Score: 4057.50 Matches: 740
 Percent Similarity: 98.58% Conservative: 21
 Best Local Similarity: 95.85% Mismatches: 10
 Query Match: 82.87% Indels: 1
 DB: 1 Gaps: 1

US-09-774-490-1 (1-2709) x SM3A_RAT (1-772)

Qy 200 ATGGCGCTGTTAACTAGGATGTCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
 Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyIleLeuLeuThrAlaArgAla 20


```

Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY 2477 AGGACCCAGCAATTTGAGAGGCGCACCCAGGAGCTGTC 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 3
SM3A_MOUSE
ID SM3A_MOUSE STANDARD; PRT; 772 AA.
AC O08665; Q62180; Q62215;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D)
DE (Sema D).
GN SEMA3A OR SEMAD OR SEMD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97470885; PubMed=9331345;
RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
RA Yagi T.;
RT "Disruption of semaphorin III/D gene causes severe abnormality in
RT peripheral nerve projection.";
RL Neuron 19:519-530(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Kimura T., Fishman M.C.;
RT "cDNA sequence of mouse collapsin/semaphorin III.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 107-772 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=95267432; PubMed=7748562;
RA Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
RA Goodman C.S., Kolodkin A.L.;
RT "Semaphorin III can function as a selective chemorepellent to pattern
RT sensory projections in the spinal cord.";
RL Neuron 14:949-959(1995).
CC -!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
CC NORMALLY TERMINATE DORSALLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
CC (E11) IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and/or commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Db EMBL; X85993; CAA59985.1; -

```

```

DR EMBL; D85028; BAA19773.1; -.
DR EMBL; I41541; AAL77611.1; -.
DR EMBL; I40484; AAA73934.1; -.
DR PIR; I48747; I48747.
DR PIR; I58169; I58169.
DR MGD; MGI:107558; Sema3a.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 772
FT DOMAIN 240 538
FT DOMAIN 579 665
FT DOMAIN 728 770
FT DISULFID 650 723
FT CARBOHYD 53 53
FT CARBOHYD 125 125
FT CARBOHYD 591 591
FT CONFLICT 193 193
FT CONFLICT 207 207
FT CONFLICT 253 253
FT CONFLICT 352 352
FT CONFLICT 403 403
FT CONFLICT 571 572
FT CONFLICT 616 620
FT CONFLICT 623 623
SQ SEQUENCE 772 AA; 88799 MW; E89A08528B10AEC3 CRC64;

Alignment Scores:
Pred. No.: 6,5e-314 Length: 772
Score: 4035.50 Matches: 737
Percent Similarity: 98.32% Conservative: 22
Best Local Similarity: 95.47% Mismatches: 12
Query Match: 82.42% Indels: 1
DB: 1 Gaps: 1

US-09-774-490-1 (1-2709) x SM3A_MOUSE (1-772)
QY 200 ATGGCGTGGTAACTAGGATGCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGACAATGTCCCAAGGCTGAAATATCTCTACAAAGAATGTTG 319
Db 21 AsnTrpAlaAsnGlyLysAsnValProArgLeuLysLeuSerTrpLysGluMetLeu 40
QY 320 GAATCCCAATGATGATCACTTCAATGGCTGGCCCAACAGCTCCAGTTATCATCCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTrpHisThrPhe 60
QY 380 CTTTGTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTrpValGlyAlaLysAspHisLeuPheSerPhe 80
QY 440 GACTGTTTATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACAGAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTrpArgArg 100
QY 500 GATCAATGCAAGTGGGCTGGAAGAAGACATCTGTAAGAAGATGCTGTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCACTTGTAGCCCTGTGGACGGGGGCTTTTTCATCCAATT 619

```

Db 121 LeuGluAlaTyraSnglnThrHisLeuTyraAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TCACCTACATTGAATTCGACATCATCTGAGGCAATATTTTAAAGTCGAGAACTCA 679
Db 141 CysThrTyriLeuGluValGlyHisProGluAspAsnIlePheLeuSnglnAspSer 160
QY 680 CATTTCGAAACGGCGTGGGAGAGTCCATATGACCTTAAGCTCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyraPheProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATTAATCTCTGAACTCGAGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyraSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTCTGGGCAACCAACCAATCAGACAGACAGATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleAlaArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGATGACATCCTGAAGT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTCTTCCTGAAATGCAATAGATGGAGAACACTCTCGAAAAGCT 979
Db 241 AspLysValTyraPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAATGACACATCTCTCAAGCTCTGCTGATTTGCTCAGTCCAGGTCCAATGCG 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTGATCAACTGAGAGTGTATTCCTAATGAACTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
QY 1160 AATCCAGTTGTATAGAGTGTATGAGTTCAGACTTCAGTAACATTTTCAAGGATCAGCGTG 1219
Db 321 AsnProIleValTyraGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATAGATGATGAGAGGTTCTTCTGTTGTTCCATATGCCACACAGGAT 1279
Db 341 CysMetTyraSerMetSerAspValArgArgValPheLeuGlyProTyraAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGTGCTTATCAAGGAAGTCCCTATCCAGGCGCAGGAAT 1339
Db 361 GlyProAsnTyraGlnTrpValProTyraGlnGlyArgValProTyraProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGATCTTCAAGAGGACTTCCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACTTTGCAAGAGTCAATCCAGCATGTATCAATCCAGTGTTCCTATGACATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyraAsnProValPheProIleAsnAsnArgPro 420
QY 1460 ATATGATCAAAACGGATGTAAATTAATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyraGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGGAAGATGATGATGTTATGTTTATCGGAACAGATGTTGGGACCTTTCTAAA 1579
Db 441 AlaGluAspGlyGlnTyraAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 CTAGTTTCATTCCTTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCGAAGATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTCCTCGGAACCGACTGCTATTTCAGCAATGGAGCTTTCCATCAAGCAGCAACAA 1699

Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGTTTCAACGGCTGGGTTCCCGAGTCCCTTTACACCGGTGATATTTAC 1759
Db 501 LeuTyriLeGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGTGCTCGCCGAGACCCCTTACTGTCTGTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyraCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCCAAGAGACGACACAGACGACAAAGATATAA 1879
Db 541 SerCysSerArgTyraPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTCTGTTTCACTTACACTTA---CACCATGATATATCACCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGlnHisHisAspAsnHisHisGlyProSer 580
QY 1937 CCTGAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTTCGAATGCAGTCCG 1996
Db 581 LeuGluGluAlaGlyIleIleTyraGlyValGluAsnSerThrPheLeuGluCysSerPro 600
QY 1997 AAGTCGAGAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGAGCGAAAGAA 2056
Db 601 LysSerGlnArgAlaLeuValTyraTrpGlnPheGlnArgArgAsnGluAspArgLysGlu 620
QY 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTCTCTCAGTAGTCTA 2116
Db 621 GluIleArgMetGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
QY 2117 CAACAGAGAGATTCAGGCAATTTACTCTGCTCCTCGGTGGAAACATGGGTTTCATACAAACT 2176
Db 641 GlnLysLysAspSerGlyAsnTyraLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
QY 2177 CTTCTTAAGCTAACCCCTGGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCATAAA 2236
Db 661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLys 680
QY 2237 GATGATGATGAGATGGCTCTTAAGACCCNAAGAAATGCTCAATAGCATCAGACCTAGCCAG 2296
Db 681 AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln 700
QY 2297 AAGTCTGCTGACAGACTTCTCATGAGCTCATCAACACCCCAATCTCAACACAGATGGAT 2356
Db 701 LysValTrpTyraArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY 2357 GAGTCTGTGAACAAGTTTGGAAAAGGACCCGAAAACAACTCGCGCAAGCCAGGACAT 2416
Db 721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHis 740
QY 2417 ACCCGAGGAACAGTAAACAAATGGAGCCTTACAGNAATTAAGAAAGGTAGAACAGG 2476
Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY 2477 AGGACCCACGAAATTTGAGAGGCGCCACCGAGGAGTGTCT 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 4
SM3A_CHICK
ID SM3A_CHICK STANDARD; PRT; 772 AA.
AC 090607;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Collapsin-1) (COLL-1).
GN SEMA3A OR COLL1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
 RC 666-680.
 RX TISSUE=Brain;
 RA MEDLINE=94006554; PubMed=8402908;
 RA Luo Y., Raible D., Raper J.A.;
 RT "Collapse: a protein in brain that induces the collapse and paralysis
 RT of neuronal growth cones.";
 RL Cell 75:1217-1227(1993).
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
 CC AND MUSCLE, MODERATE LEVELS IN LUNG, BURSA, AND HEART AND
 CC VIRTUALLY ABSENT IN LIVER. COLLAPSEIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U02528; AAC59638.1; --
 DR PIR; A49069; A49069.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 772 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 576 665 IG-LIKE C2-TYPE.
 FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 650 723 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 772 AA; 88867 MW; E91E09DE0CC940AC CRC64;
 Alignment Scores:
 Pred. No.: 6 91e-294 Length: 772
 Score: 3784.50 Matches: 683
 Percent Similarity: 95.34% Conservative: 53
 Best Local Similarity: 88.47% Mismatches: 35
 Query Match: 77.30% Indels: 1
 DB: 1 Gaps: 1
 US-09-774-490-1 (1-2709) x SM3A_CHECK (1-772)
 QY 200 ATGGGCTGTTAACTAGGATTCGTCTTTCTGGGAGTATTACTTACACAGACGCA 259
 Db 1 MetGlyTrpLeuArgGlyIleAlaLeuLeuSerLeuGlyValLeuAlaGlyArgVal 20
 QY 260 AACTATCAGAAATGGGAGACAACTGTGCCAAGCTGAATATCTTACAAAGATTTG 319
 Db 21 AsnCysGlnHisValIysAsnValProArgLeuLeuLysSerTyrIysGluMetLeu 40

QY 320 GAATCCAAATGTGATCACTTTCAATCGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
 Db 41 GluSerAsnAsnIleValAsnPheAsnGlyLeuAlaAsnSerSerSerfYHisThrPhe 60
 QY 380 CTTTGGATGAGGACGAGTAGCTGTATCTTGGAGCAAGGATCATATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGA 499
 Db 81 AsnLeuValAsnIleLysGluTyrGlnLysIleValTrpProValSerHisSerArgArg 100
 QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGTAATTTTCATCAAGTA 559
 Db 101 AspGluCysLysTyrAlaGlyLysAspIleLeuArgGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGATATATACAGACTCATTGTGACGCTGTGGAACGGGGCTTTTCATCAATT 619
 Db 121 LeuLysThrTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProMet 140
 QY 620 TGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGAGACTCA 679
 Db 141 CysThrTyrIleGluValGlySerHisProGluAspAsnIlePheArgMetGluAspSer 160
 QY 680 CATTTTGAAAAACGGCTGGGAAGAGTCCATATGACCCCTAAAGCTCTGACAGCATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTTACTCTGGAACTCGAGCTGATTTTATCGGGGAGACTTTGCT 799
 Db 181 LeuValAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAACTCTTGGGCACACCACCAATCAGGACAGACAGCAGCATGATTTCCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProLeuArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCCAAAGTTTCATTAGTGCCTCCACCTCATCTCAGAGAGTGACAAATCCTGAAGAT 919
 Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuLeuLeuProGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGAGAGAACACACTCTGGAAGGCT 979
 Db 241 AspLysIleTyrPhePheArgGluAsnAlaIleAspGlyGluHisThrGlyLysAla 260
 QY 980 ACTCAGCTAGATAGGTGCAGATATGCAAGATGACTTTGGAGGSCACAGAGTCTGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACATTTCTCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCCAATGCG 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACATCAATTTGATGAACCTGAGGATGTATTCCTTAATGAATTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
 QY 1160 AATCCAGTTGTATATGGAGTCTTACGACTTCCAGTACATTTTCAAGGGATCACCGCTG 1219
 Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTCTTCTGCTTCCATATGCCACAGGAT 1279
 Db 341 CysMetTyrSerMetThrAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCCACTATCAATGGGTGCTTTATCAAGAGAGAGTCCCTATTCACGGCCAGGAAC 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCACGACAAATTTGGTGGTTTGACCTCAAGAGGACCTTCCCTGATGATGCTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspGluValIle 400

QY 1400 ACCTTTGCAAGAGTCCATCCAGCCATGTAATCAATCCAGTGTTCCTTATGAAACAATCGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnSerArgPro 420
 QY 1460 ATAGTGATCAAAACCGATGAATTAATATCAAAATTTACACAAATTTGTCGTAGACCGAGTGGAT 1519
 Db 421 IleMetIleLysThrAspValAspTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGGACAGTATGATGTTATGTTATCGGACAGATGTTGGACCGCTTCTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspIleGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTTAAGCAGACTGCTGATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
 Db 461 ValValSerIleProLysGlnThrPheHisGluLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTTCACATGAAGCAGCAACAA 1699
 Db 481 ThrValPheArgGluProThrValIleSerAlaMetLysIleSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTTGTTCAACGCTGGGTTCGCCAGCTCCCTTTACACGGGTGCTATTTTAC 1759
 Db 501 LeuTyrIleGlySerAlaThrGlyValSerGlnLeuProLeuHisArgCysAspValTyr 520
 QY 1760 GCGAAGCGTGTGCTGAGTGTTCCTCCGCGCAGACCTTACTGTGCTTGGATGTTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysLysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGACGCGACAGCAAGCAAGATATAAAGAT 1879
 Db 541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
 QY 1880 GGAGACCCACTGACTCAGCTGTTTCAGACTTA---CACCATGATAATCACCATGCCACAGC 1936
 Db 561 GlyAspProLeuThrHisCysSerAspLeuGlnHisHisAspAsnProSerGlyGlnThr 580
 QY 1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAGACATTTTGGAAATCGAGTCG 1996
 Db 581 LeuGluGluLysIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
 QY 1997 AAGTCGCAGAGCGCTGCTTATTTGGCAATTCAGAGCGCAAAATCAAGAGCGCAAAAGAA 2056
 Db 601 LysSerGlnArgAlaIleValTyrTrpGlnPheGlnLysGlnAsnAspHisLysVal 620
 QY 2057 GAGATCAGAGTGATGATCATATCATCAGCAGACAGATCAAGGCTCTGCTAGCTAGTCTA 2116
 Db 621 GluIleLysValAspAspArgMetIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
 QY 2117 CAACAGAGCATTCAGGCAATACCTCTGCCATGCGTGGACATGGGTTTCATACAACT 2176
 Db 641 GlnArgArgAspSerGlyIleTyrPheCysHisAlaValGluHisGlyPheIleGlnThr 660
 QY 2177 CTTCTTAAGGTAACCTCGAAGTCATTTGACAGACGATTTTGGAAAGACTTCTTCAAAA 2236
 Db 661 LeuLeuLysValThrLeuGluValIleAspThrAspHisLeuGluGluLeuHisLys 680
 QY 2237 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCATAGATGATGACCTAGCCAG 2296
 Db 681 GluGluAspAlaSerLysThrLysAspAlaThrAsnSerMetThrProSerGln 700
 QY 2297 AAGGTCGTGTACAGACTTCATGAGCTCATCAACCCCACTCAACACGATGGAT 2356
 Db 701 LysIleTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
 QY 2357 GAGTTCGTGAACAAGTTTGGAAAAGGGAACGAAACAAACGTCGCGCAAGCCAGACAT 2416
 Db 721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProIleAsn 740
 QY 2417 ACCCAGGGAACAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGTGAACAGC 2476
 Db 741 AlaGlnValAsnThrAsnLysTrpLysHisLeuGlnGlnAsnLysLysGlyArgAsnArg 760
 QY 2477 AGGACCCACCAATTTGAGAGGCGCACCCAGGAGTGTC 2512

Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772
 |||||

RESULT 5

SZ1B_BRARE
 ID SZ1B_BRARE STANDARD; PRT: 778 AA.
 AC Q9W686;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin z1b precursor (Semaphorin 1B) (Sema-z1b).
 GN SEMAZ1B OR SEMAZ3AB.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425174; PubMed=10495275;
 RA Roos M., Schachner M., Bernhardt R.R.;
 RT "Zebrafish semaphorin z1b inhibits growing motor axons in vivo.";
 RL Mech. Dev. 87:103-117(1999).
 CC -!- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
 CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
 CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
 CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
 CC VENTRALLY EXTENDING MOTOR AXONS.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; AF083382; AAD28103.1; -
 DR ZFIN; ZDB-GENE-991209-6; sema3ab.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig_1.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 778
 FT SEMAPHORIN Z1B.
 FT DOMAIN 241 539
 FT IG-LIKE C2-TYPE.
 FT DOMAIN 579 668
 FT ARG/LYS-RICH (BASIC).
 FT DOMAIN 721 776
 FT DISULFID 652 716
 FT CARBOHYD 54 54
 FT CARBOHYD 127 127
 FT CARBOHYD 593 593
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 778 AA; 88904 MW; 4D36F4323AE21895 CRC64;
 Alignment Scores:
 Pred. No.: 749e-248 Length: 778
 Score: 3207.50 Matches: 591

320	ProLysAsnProIleIleTyrAlaValPheThrThrSerSerAsnIlePheLysGlySer	339
1214	GCCGTGTGATGATAGCATGAGTGATGAGAGGGTGTTCTTGCTCCATATGCCAC	1273
340	AlaValCysMetTyrSerMetAlaAspIleIleArgArgValPheLeuGlyProTyrAlaHis	359
1274	AGGATCGACCCCACTAATCGGTGCTTTATCAAGGAAGAGTCCCTATCCACGGCCA	1333
360	ArgAspGlyProAsnTyrGlnTrpValProPheLeuAsnArgValProTyrProArgPro	379
1334	GGAACTTGTCCCGCAAAACATTTGGTGGTTTGCACCTCTCAAGAAGACCTTCCTGAT	1393
380	GlyThrCysProSerLysThrPheAspGlyPheGluSerThrLysAspPheProAspAsp	399
1394	GTATATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAAT	1453
400	ValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsn	419
1454	CGCCAAATAGTGATCAAAACCGGATGTAAATTATCAAAATTTACAAATGTGCTAGACCGA	1513
420	HisProIleIleLysThrAspValAspTyrGlnPheThrGlnIleValValAspArg	439
1514	GTGGATCGCAAGATGACAGTAGTAGTGTATGTTTATCGAACAGATGTTGGGACGGTT	1573
440	ValGluAlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspMetGlyThrVal	459
1574	CTTAAAGTAGTTTCAATTCCTAAGGAGACTGTGTATGATTATAGAAGAGGTTCTGCTGAA	1633
460	LeuLysValValSerIleProArgGlyThrTrpHisAspLeuGluGluValLeuLeuGlu	479
1634	GAATGACAGTTTTTCGGGAACCGACTGCTATTTCACGAATGGAGCTTTCACATAAGCAG	1693
480	GluMetThrValPheArgGluProThrAlaIleThrAlaMetGluLeuSerThrLysGln	499
1694	CAACAACATATATTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGAT	1753
500	GlnGlnLeuTyrLeuGlySerAlaIleGlyValSerGlnMetProLeuHisArgCysAsp	519
1754	ATTTCACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTGCTGGAT	1813
520	ValTyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAsp	539
1814	GGTTCGCACTGTTCTCGCTATTTTCCCACTGCAAGAGACGCAACAGACAGATATA	1873
540	GlySerGlnCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIle	559
1874	AGAAATGGACACCCACTGACTCTACTGTTACAGACTTA---CACCATGATATACCATGGC	1930
560	ArgAsnGlyAspProLeuThrGlnCysSerAspLeuGlnHisHisAspGluAlaAspGly	579
1931	CACAGCCCT--GAAGAGAGAATCATCTATGGTGTAGAGAAATAGTAGACATTTTGGAA	1987
580	GluAlaGlyLeuLeuAspLysThrValTyrGlyValGluAsnSerSerPheLeuGlu	599
1988	TGCAGTCCGAAGTCGCAGAGAGCGCTGGTCTATTGGCAATTCAGAGCGGCAATGAAGAC	2047
600	CysSerProLysSerGlnArgAlaLeuIleTyrTrpGlnPheGlnArgHisGlyGluAsp	619
2048	CGAAAGAAGAGATCAGATGGATGATCATATCATCAGGACAGATCAAGCGCTTCGTCTA	2107
620	HisLysLeuGluIleLysSerAspGluArgValLeuGlyThrGluGlnGlyLeuLeuIle	639
2108	CGTAGTCTACAAACAGAGGATTCACGAATTAACCTCTGCCATCGCTGGACATGGGTTC	2167
640	ArgSerLeuHisGlnLysAspSerGlyValTyrTyrCysHisAlaValGluHisGlyPhe	659
2168	ATACAAACTCTTCTTAAGGTAAACCTCGGAAGTCATTGACACAGAGCATTTTGAAGAAC	2227
660	IleGlnThrLeuLeuArgLeuThrLeuAsnValIleProAlaGluHisLeuAspAspLeu	679
2228	CTTCATAAAGATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAAATAGCATGACA	2287
680	LeuHisArgAspPro-----ProAspThrAsnAspProAlaAsnGly-----	693

CC NEURONAL POPULATIONS. BINDS TO NEUROPILIN (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THROF OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AC004957; AAC83081.1; ALT_SEQ.
 CC Genew; HGNC:10726; SEMA3D.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00447; IG; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS00835; IG_LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 39
 CC CHAIN 40 777
 CC DOMAIN 257 555
 CC DOMAIN 592 680
 CC DOMAIN 739 773
 CC DISULFID 665 731
 CC CARBOHYD 139 139
 CC CARBOHYD 607 607
 CC CARBOHYD 724 724
 CC SEQUENCE 777 AA; 89651 MW; 3F7B0D7AF50F53BD CRC64;
 FT
 FT CHAIN 40 777
 FT DOMAIN 257 555
 FT DOMAIN 592 680
 FT DOMAIN 739 773
 FT DISULFID 665 731
 FT CARBOHYD 139 139
 FT CARBOHYD 607 607
 FT CARBOHYD 724 724
 SQ SEQUENCE 777 AA; 89651 MW; 3F7B0D7AF50F53BD CRC64;
 Alignment Scores:
 Pred. No.: 2,92e-171 Length: 777
 Score: 2247.50 Matches: 414
 Percent Similarity: 73.94% Conservative: 142
 Best Local Similarity: 55.05% Mismatches: 175
 Query Match: 45.90% Indels: 21
 DB: 1 Gaps: 10
 US-09-774-490-1 (1-2709) x SM3D_HUMAN (1-777)
 QY 275 AAGAACAAATGTGCCAAGGCTGAAATTCCTCAAGAAATGTGGAATCCAAACAAATGTG 334
 Db 39 LysGlnAsnIleProArgLeuLysLeuThyTyLysAspLeuLeuSerAsnSerCys 58
 QY 335 ATCACTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTCTCTTTTGGATGGGAA 394
 Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
 QY 395 CGAGTAGGCTGTATGCTGGAGCAAGGATCATATTTTCATTCGACCTGGTTAATATC 454
 Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 QY 455 ---AAGGATTTTCAAGATTTGTGCGCCAGTATCTTACCCAGAGAGAGATGAATGCAAG 511
 Db 99 AsnLysAsnPhelLysLysIleTyTrpProAlaAlaLysGluArgValGluLeuCysLys 118
 QY 512 TGGCGTGGAAAGACATCTCGTGAAGATGTGCTAATTTTCATCAAGTACTTAAGGCATAT 571
 Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPhelIleArgValLeuGlnProTyTr 138
 QY 572 AATCAGACTCACTTGTACGCTGTGGAACGGGGGCTTTTTCATCCAAATTTGCACCTACATT 631

Db 139 AsnLysThrHisIleTyValCysGlyThrGlyAlaPheHisProIleCysGlyTyIle 158
 QY 632 GAATTTGACATATCTCTGAGCAACAATATTTTAACTGGGAACTCACATTTTGAAC 691
 Db 159 AspLeuGlyValTyLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
 QY 692 GGCCTGGGAAGAGTCCATATGACCTAAGCTGTCACAGCATCCCTTTTAAATAGATGA 751
 Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 QY 752 GAATTTACTCTGGAAGTGCAGTCAATTTTATGGGCGAGACTTTGTCTTCTTCCGA 811
 Db 199 TyrLeuTySerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
 QY 812 CTTGGG-----CACCACCACCATCAGGACAGCAGCATGATCCAGGTGGCTC 862
 Db 219 LeuGlyProThrHisAspHisIleTyTrpIleArgThrAspIleSerGluHisTyTrpLeu 238
 QY 863 AATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGATGAC 922
 Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyAsnProAspAsp 258
 QY 923 AAAGTATATCTTTTCTTCCTGTAATAATGCAATAGATGGAGAACACTCTGGAAAGCTACT 982
 Db 259 LysIleTyPhePhePheArgLysSerSerGlnGluGlySerThrSerAspLysThrIle 278
 QY 983 CACCTAGATAGTGCAGATATGCAAGATGACATTTTGAGGGCGACAGAGTCTGTGTAAT 1042
 Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
 QY 1043 AAATGGACAACATTCCTCAAGCTCGTGTGATTTGCTCAGTGCCAGGTGCCAAATGGCA 1102
 Db 299 LysTrpThrThrPheLeuLysAlaArgLeuLeuCysSerIleProGlySerAspGlyAla 318
 QY 1103 GACACTCAATTTGATGAACCTCAGAGATGTATCTCTAATGAATTTTAAAGATCTTAAAT 1162
 Db 319 AspThrTyPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAspGluArgAsn 338
 QY 1163 CGAGTTGTATGAGTGTATTCAGCTTCCAGTAACTTTTCAGGGATCAGCCGTGTGT 1222
 Db 339 ProValValTyGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
 QY 1223 ATGTATAGCATGAGTGTGAGAGGCTGTCTTGTGCTCATATGCCACAGGATGGA 1282
 Db 359 ValTySerMetAlaAspIleArgAlaValPheAsnGlyProTyAlaHisLysGluSer 378
 QY 1283 CCCAACTATCAATGGTGCCTTATCAAGAGAGAGTCCCTTATCCACGGCCAGGAATGT 1342
 Db 379 AlaAspHisArgTrpValGlnTyArgGlyArgIleProTyTrpProArgProGlyThrCys 398
 QY 1343 CCCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTCTGATCATGTATA 1399
 Db 399 ProSerLysThrTyAspProLeuLysSerThrArgAspPheProAspAspValIle 418
 QY 1400 ACCTTTGCAAGAGTCAATCCAGCCATGACCAATCCAGTGTCTTCTATGAACAAATCCCCA 1459
 Db 419 SerPheIleLysArgHisSerValMetTyLysSerValTyProValAlaGlyGlyPro 438
 QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAATTTCTGTAGACCGATGGAT 1519
 Db 439 ThrPheLysArgIleAsnValAspTyArgLeuThrGlnIleValValAspHisValIle 458
 QY 1520 CGAGAGATGACATGATGTTATCGGAACAGATGTTGGGACCGCTTCTTAAA 1579
 Db 459 AlaGluAspGlyGlnTyAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCTGTGGAAGAAATG 1639
 Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluLeu 497
 QY 1640 ACAGTTTTCGGGAACCGAGCTCTTATTCAGCAATGAGGAGCTTTCACCTAAGCAGCA 1699

Db 498 GlnIlePheLysHisSerSerIleIleLeuLeuMetGluLeuSerLeuLysGlnGln 517
 QY 1700 TATATATTGTTTCAACGGCTGGGTGTCACGCTCCCTTTACACCGGTGATATTTAC 1759
 Db 518 LeuTyrlleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
 QY 1760 GGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1819
 Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
 QY 1820 GCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1879
 Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
 QY 1880 GGAGACCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
 Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
 QY 1940 GAAGAGAGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999
 Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
 QY 2000 TCGCAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
 Db 617 SerGlnGlnAlaThrIleLysTyrTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
 QY 2060 ATCAGAGGATGATCATATCATCATCAGCAGCATCAAGCGCTTCTGCTAGCTAGCTACAA 2119
 Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
 QY 2120 CAGAGGATTCAGCAATTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2179
 Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
 QY 2180 CTTAAGGTAACTTGAAGTTCATTCATCAGCAGCATCAAGCGCTTCTGCTAGCTAGCTACAA 2239
 Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
 QY 2240 GATGATGAGATGCTTCTGAGCAACCAAGAAATGCTCAATGATGATGATGATGATGATG 2299
 Db 697 His-----GluGluGlyLysValLysAsp-----LeuLeuAlaGluSerArg 710
 QY 2300 GTCTGTTACAGACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2359
 Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
 QY 2360 TTCTGTGAACAGTTTGGAAAGGAGCCGAAACCAACGTCGCGCAAGCCAGGACATACC 2419
 Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
 QY 2420 CCAGGGAACAGTAACAATGGAAGCATTCAAGAAATAGAAAGGTAGAAACAGGAGG 2479
 Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysLysLysLysLysLysLys 763
 QY 2480 ACCAC---GAATTGAGAGGCGCCCGAGGAGTGTC 2512
 Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 8

SM3D_CHICK
 ID SM3D_CHICK STANDARD; PRT; 761 AA.
 AC Q90663;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Senaphorin 3D precursor (Collapsin-2) (COLL-2).
 GN SENAP3D OR COLL2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]

US-09-774-490-1 (1-2709) x SM3D_CHICK (1-761)
 QY 188 AGCGTCTGCGAGCATGGCTGTTAACTAGGATGTTCTGCTTTCTGGGAGTATTACTT 247
 Db 8 AshAlaCysSer-----LeuLeuSerLeuAlaMetLeuPheProValThrGly 24
 QY 248 ACAGCAAGACAACTATCATGATGGAAGCAACATGTCGCAAGCTGAAATTTATCTTAC 307
 Db 25 ThrSer-----LysGlnAsnIleProArgLeuLysLeuSerTyr 37
 QY 308 AAAGAATGTTGAATCCAAATGTGATGATCACTTTCAATGGCTTGGCCACAGCTCCAGT 367
 Db 38 LysAspLeuLeuSerAsnSerCysIleProPheLeuGlySerThrGluGlyLeuAsp 57
 QY 368 TATCATACCTTCCTTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCAC 427

Alignment Scores:
 Pred. No.: 4, 59e-171 Length: 761
 Score: 2245.00 Matches: 414
 Percent Similarity: 72.72% Conservative: 151
 Best Local Similarity: 53.28% Mismatches: 182
 Query Match: 45.85% Indels: 30
 DB: 1 Gaps: 12

SEQUENCE FROM N.A.
 MEDLINE=95329269; PubMed=7605628;
 RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
 RT "A family of molecules related to collapsin in the embryonic chick nervous system";
 RL Neuron 14:1131-1140(1995).
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: DEVELOPING SPINAL CORD AND DEVELOPING VISUAL SYSTEM. COLLAPSIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch)

EMBL: U28240; AAA86896.1; --
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003599; Ig.
 InterPro; IPR003006; Ig_MHC.
 InterPro; IPR003659; Plexin-like.
 InterPro; IPR001627; Sema.
 Pfam: PF01403; Sema; 1.
 SMART; SM00409; IG; 1.
 SMART; SM00423; PSI; 1.
 SMART; SM00630; Sema; 1.
 SMART; SM00630; Sema; 1.
 Signal; PS00835; IG_Like; 1.
 Signal; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 761 SEMAPHORIN 3D.
 FT DOMAIN 245 543 SEMA.
 FT DOMAIN 552 670 IG-LIKE C2-TYPE.
 FT DOMAIN 727 757 ARG/LYS-RICH (BASIC).
 FT DISULFID 653 719 BY SIMILARITY.
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 761 AA; 87300 MW; 3E09AE3DBA53F46B CRC64;

AC Q9W6G6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin Z2 precursor (Semaphorin 2) (Sema-Z2).
 GN SEMAZ2 OR SEMA2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99112778; PubMed=9915572;
 RA Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Raper J.A.,
 RA Kuwada J.Y.;
 RT "Analysis of a Zebrafish semaphorin reveals potential functions in
 RT vivo";
 RL Dev. Dyn. 214:13-25(1999).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE GUIDANCE OF SEVERAL AXON
 CC PATHWAYS.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN A DYNAMIC AND RESTRICTED PATTERN
 CC DURING THE PERIOD OF AXON OUTGROWTH.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF124485; RAD21310.1; --
 CC ZFIN; ZDB-GENE-990715-2; sema2.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 41
 CC CHAIN 42 784 SEMAPHORIN Z2.
 CC DOMAIN 261 559 SEMA.
 CC DOMAIN 661 740
 CC DOMAIN 741 762 ARG/LYS-RICH (BASIC).
 CC DOMAIN 768 733 BY SIMILARITY.
 CC DISULFID 668 733
 CC CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 490 490 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SEQUENCE 764 AA; 87859 MW; A3BD95C2C479D7AE CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 1-72e-161 Length: 764
 Score: 2125.00 Matches: 390
 Percent Similarity: 71.37% Conservative: 156
 Best Local Similarity: 50.98% Mismatches: 191
 Query Match: 43.40% Indels: 28
 DB: 1 Gaps: 11
 US-09-774-490-1 (1-2709) x SMZ2_BRARE (1-764)
 QY 215 AGGATTGCTGCTTTCTGCGGAGTATTACTTACGACGAGCAAACTATCAGAATGGG 274
 Db 20 ArgPheSerCysAlaTrpTrpSerThrSerValMetPhePheSerLeuProGluGly 39
 QY 275 -----AGAACAAATGTGCCAGGCTGAAATTTATCTCTACAAAGAAATGTTGAATCC 325

Db 40 AsnCysMetLysGluSerLeuProArgValLysLeuGlyTyrLysAspLeuIleHisSer 59
 QY 326 AACAAATGTGATCACTCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCACTCTCTTTTG 385
 Db 60 ArgSerValProPheThrGlySerSerGluGlyGlnHisPheGlnThrValLeuLeu 79
 QY 386 GATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTTCATTCGACCTG 445
 Db 80 AspGluGluArgSerArgLysLeuLeuGlyAlaLysAspHisValTyrLeuLeuAspPro 99
 QY 446 GTTAAATATC---AAGGATTTTCAAAAGATTGTGTCGCCAGTATCTTACACGAGAGAT 502
 Db 100 AspAsnIleAsnLysHisProLysLysLeuSerTrpProAlaSerArgAspArgValGlu 119
 QY 503 GAATCAAGTGGCTGGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCATCAAGTACTT 562
 Db 120 MetCysIleLeuAlaGlyLysAsnProLeuThrGluCysAlaAsnPheIleArgValLeu 139
 QY 563 AAGCATATAATCAGACTCCTGTGACGCTGTGGAACGGGGCTTTTCATCCAAATTTGC 622
 Db 140 HisSerTyrAsnArgThrHisValTyrAlaCysGlyThrGlyAlaPheHisProThrCys 159
 QY 623 ACCTACATTGAATTTGGACATCATCTCGAGACAATATTTTAACTGGAGAACTCATCAT 682
 Db 160 AlaPheLeuGluIleLysGlyHisLysGluAspArgTrpLeuLeuLeuHisSerAsnThr 179
 QY 683 TTTGAAAACGGCTGGGGAAGAGTCCATATGACCCCTAAGCTGCTCAGACGATCCCTTTTA 742
 Db 180 MetGluSerGlyArgMetLysCysProPheAspProAsnGlnProPheAlaSerValLeu 199
 QY 743 ATAGATGAGAAATATATCTCTGGAACCTGAGTATTTTATGGGCGGAGACTTTGCTATC 802
 Db 200 ThrAspGlnTyrLeuTyrAlaGlyThrAlaSerAspPheLeuGlyLysAspSerThrPhe 219
 QY 803 TTCGAACTCTTGGG-----CACCACACCCCAATCAGACGACGAGCATGATCTCC 853
 Db 220 ThrArgSerLeuGlyProProHisGlnGlnTyrIleArgThrAspIleSerGluAsp 239
 QY 854 AGTGGCTCAATGATCAAAAGTTCTATTAGTGCACCTCATCTCAGAGAGTGACATCTCT 913
 Db 240 TyrTrpIleAsnGluGlyLysPheIleSerAlaHisProIleSerAspThrTrpAsnPro 259
 QY 914 GAAGATGACAAATATATCTTTCTCGTGAAATGCAATAGATGAGAGAACACTCTCGA 973
 Db 260 AspAspAspLysIleTyrPhePhePheArgGluAlaSerArgAspGlySerThrThrAsp 279
 QY 974 AAAGCTACTCAGCTAGATAGTGCAGATATGCAAGATGATCTTTGGAGGGCAGCAAGCT 1033
 Db 280 LysSerValLeuSerArgValAlaArgIleCysArgAsnAspValGlyLeuArgSer 299
 QY 1034 CTGGTGAATAAATGACAAACATCTCTCAAGCTCGTCTGATTGCTCAGTGCAGGTCCA 1093
 Db 300 LeuThrAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerIleProGlyPro 319
 QY 1094 AATGCATTGACACTCATTTTGTGCACTGAGAGTGTATTCCTTAATGAACCTTTAAAGAT 1153
 Db 320 AspGlyValAspThrHisPheAspGluLeuGlnAspIlePheLeuLeuProSerArgAsp 339
 QY 1154 CCTAAATTCAGTTGTATGAGAGTGTTCAGACTTCCAGTCCAGTAACATTTTCAGGATCA 1213
 Db 340 GluLysAsnProMetValTyrGlyValPheThrThrSerSerIlePheLysGlySer 359
 QY 1214 GCGGTGTGTATGATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCAC 1273
 Db 360 AlaValCysValTyrThrMetGluAspIleArgAlaAlaPheAsnGlyProTyrAlaHis 379
 QY 1274 AGGATGACCACTATCAATGGTGCCTTATCAAGAGAGAGTCCCTATCCACGGCA 1333
 Db 380 LysGluGlyProAspHisArgTrpValGluTyrGluGlyArgGlyIleProTyrProArgPro 399
 QY 1334 GGAATCTGTCCACGAAAAACATTTGGT---GGTTTGTACTCTACAAAGAGACTTCTCTGAT 1390

Db 400 GlyThrCysProSerArgThrTyrAspProHisIleLeuThrThrLysAspPheProAsp 419
QY 1391 GATGTTATACCTTTTCAGAGATCATCCAGCATCAATCAATCCAGTGTTCCTATGAC 1450
Db 420 GluValIleSerPheIleArgLeuHisProLeuMetTyrGlnSerValHisProMetThr 439
QY 1451 AATCGCCCATAGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTGAGAC 1510
Db 440 GlyArgProIlePheThrArgIleAsnThrGluTyrArgLeuThrGlnIleIleValAsp 459
QY 1511 CGAGTGGATCGAAGATCGACAGTATGATGTTATGTTATCGGACAGATGTTGGACC 1570
Db 460 ArgValAlaAlaGluAspGlyGlnTyrAlaValMetPheLeuGlyThrAspMetGlySer 479
QY 1571 GTTCTTAAAGTAGTTTCAATCTCAAGGACATGTTGATGATTTAGAGAGTTCCTGCTG 1630
Db 480 ValLeuLysValValSerIleThrGlnGluAsnTrp---SerSerGluGluIleIleLeu 498
QY 1631 GAAGAATACACAGTTCCTCGGAACCGACTGCTGTTATTCAGCAATGAGCTTCCACTAAG 1690
Db 499 GluGluLeuGlnValPheLysAsnProSerProIleLeuAsnMetGluValSerSerLys 518
QY 1691 CAGCAACAATATATGTTTCAACGGTGGGTTGCCAGCTCCCTTTACACCGTGT 1750
Db 519 GlnGlnGlnLeuPheValGlySerAspGlyLeuValGlnValSerLeuHisArgCys 538
QY 1751 GATATTACGGAAACGGTGTCTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1810
Db 539 GlnIleTyrGlyGlnGlyCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrp 558
QY 1811 GATGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1870
Db 559 AspGlyThrGlnCysSerArgTyrIleProAlaSerLysArgAlaArgGlnAsp 578
QY 1871 ATAAGAATGGAGACCATCATCTACTGTTACAGCTTACCATCATCAATCACCATGCG 1930
Db 579 IleLysHisGlyAspProSerHisCysTrpAspThr-----GluAspValLeuGly 596
QY 1931 CACAGCCCTGAGAGAGATCATCTATGTTGAGAAATAGTAGACATTTTGGAAATGC 1990
Db 597 ArgAsnValGluGlnLysValLeuTyrGlyValGluSerAsnSerPheLeuGluCys 616
QY 1991 AGTCCGAATCGCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2050
Db 617 ValSerLysSerGlnGlnAlaLeuIleArgTyrPheValLeuLysProGlyValAspHis 636
QY 2051 AAGAAGATCATGATGATCATATCATCAGGACAGATCAAGCGCTTCTGTACGT 2110
Db 637 ArgGlnGluIleLysProAspGluArgValLeuIleThrAspArgGlyLeuLeuIleArg 656
QY 2111 AGTCTACACAGAGATTCAGGCATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170
Db 657 TrpLeuGlnArgGlyAspAlaGlySerTyrPheCysThrSerGlnGlnHisArgPheThr 676
QY 2171 CAAACTCTTCTTAAGTCAACCTGGAAGTCATTGACACAGAGATTTGGAAGAACTTCT 2230
Db 677 ArgThrLeuLeuHisValSerLeuHisIleLeuAspArgGlyGlnIleAsn-----Ala 694
QY 2231 CATAAGATGATGATGATGCTCTAAGACAAAGAAATGTCCAATGATGATGACCT 2290
Db 695 HisGlnProAlaIleArgGluSerSerGluAsnProAlaValThrGlu-----Pro 711
QY 2291 AGCCAGAAGTCTGTGATCAGAGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2350
Db 712 ArgGlnArg-----TyrLysAspTyrLeuArgMetLysSerGlyPro---AlaArgSer 728
QY 2351 ATGGATGATGCTGTGAAACAAGTTTGGAAAGGACCGCAAAACACATCGCGCAAGGCCA 2410
Db 729 LeuAspGluTyrCysGluThrMetTyrHisArgGluLysGlnLysGlnLys-----746
QY 2411 GGACATACCCCGGGAACAGTAAACAATGAGACATTAACAAGAAATGAAGAAAGTACA 2470
Db 747 -----GlyLysTrpLysHisValGlnGlnLeuArgLysSerArg 759

QY 2471 AACAGGAGGCCAC 2485
Db 760 AsnArgArgHisHis 764
RESULT 10
SM3B HUMAN
ID SM3B HUMAN STANDARD; PRT; 749 AA.
AC Q13214; Q8TDV7; Q93018;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin V) (Sema V) (Sema A(V)).
GN SEMA3B OR SEMA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
cancer deletion region and demonstrate distinct expression patterns.";
Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RX Dante M., Wamsley P.;
RT Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RX Koyama N.;
RT Semaphorin 3B (SEMA3B) cDNA";
RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RL [4]
RP VARIANTS NSCLC CYS-348; HIS-397 AND ILE-415.
RX MEDLINE=20535986; PubMed=11085536;
RA Lerman M.I., Minna J.D.;
RT "The 630-Kb lung cancer homozygous deletion region on human chromosome
3p21.3: identification and evaluation of the resident candidate tumor
suppressor genes.";
Cancer Res. 60:6116-6133(2000).
RL [5]
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
ENDOPLASMIC RETICULUM.
CC -!- TISSUE SPECIFICITY: Expressed abundantly but differentially in a
variety of neural and nonneural tissues.
CC -!- DISEASE: Defects in SEMA3B are found in non-small cell lung cancer
(NSCLC) cell lines.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; U28369; AAD09138.1; -;
CC EMBL; U73167; AAC02731.1; -;
CC EMBL; AB083186; BAB88870.1; -;
CC PIR; G01856; G01856.
CC Genew; HGNC:10724; SEMA3B.
CC MIM; 601281; -;
CC GO; GO:0005783; C:endoplasmic reticulum; TAS.
CC GO; GO:0007411; P:axon guidance; TAS.
CC GO; GO:0007267; P:cell-cell signaling; TAS.

DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Disease mutation.
 FT SIGNAL 1 24
 FT CHAIN 25 749
 FT DOMAIN 55 497
 FT DOMAIN 573 659
 FT DOMAIN 699 702
 FT DOMAIN 724 744
 FT DISULFID 644 710
 FT CARBOHYD 82 82
 FT CARBOHYD 124 124
 FT CARBOHYD 427 427
 FT VARIANT 348 348
 FT VARIANT 397 397
 FT VARIANT 415 415
 FT CONFLICT 29 36
 FT CONFLICT 308 308
 SQ SEQUENCE 749 AA; 83121 MW; 1F3B8F63F59444F3 CRC64;

 Alignment Scores:
 Pred. No.: 6.2e-161 Length: 749
 Score: 2118.00 Matches: 404
 Percent Similarity: 67.2% Conservative: 123
 Best Local Similarity: 51.53% Mismatches: 185
 Query Match: 43.26% Indels: 72
 DB: 1 Gaps: 12

 US-09-774-490-1 (1-2709) x SM3B_HUMAN (1-749)
 QY 221 GTCCTCTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAGATGGGAAGAAC 280
 Db 13 LeuAlaLeuLeuTrpAlaValGlyLeuGlySerAlaAla-----Pro 26
 QY 281 AATGTCGAAGGCTGAATATCTTACAAAGAAATGTTGGATCCAAACATGTGATCACT 340
 Db 27 SerProProArgLeuArgLeuSerPheGlnGluLeuGlnAlaTrpHisGlyLeuGlnThr 46
 QY 341 TTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTGGATGAGGAACGGAGT 400
 Db 47 PheSer---LeuGluArgThrCysCysTyrGlnAlaLeuLeuValAspGluGluArgGly 65
 QY 401 AGCGTGATGTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATC---AAG 457
 Db 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuLeuAsnLeuAspAsnIleSerLys 85
 QY 458 GATTTTCAAGAGATTGTGGCCAGTATCTTACACCAGAAGAGATGAATGCAATGGGCT 517
 Db 86 ArgAlaLysLeuLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
 QY 518 GGAAAGACATCTCTCAAGAGATGTCTTAATTTTCATCAAGGTACTTTAAGGCATATAATCAG 577
 Db 106 GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaTyrAsnArg 125
 QY 578 ACTCACTGTAGCGTGTGGAAAGGGGGCTTTTCATCCAAATTTGACCTACATTTGAAAT 637
 Db 126 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal 145
 QY 638 GGACATCATCTCGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGT 697
 Db 145

Db 146 GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys 165
 QY 698 GGGAGAGTCCATATGACCTAAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATTA 757
 Db 166 GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValGlyGluGluLeu 185
 QY 758 TACTCTGGAACCTGACAGCTGATTTTGGGGGAGACTTTTGTCTATCTTCCGAACCTCTGGG 817
 Db 186 TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly 205
 QY 818 CACCACCAACCAATCAGGACAGCAGCATGATTCACAGTGGCTCAATGATCCAAAGTTC 877
 Db 206 GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGluProLysPhe 225
 QY 878 ATTAGTCCCACTCATCTCAGAGAGTGCATCTGAAGATGACAAATGATGACAAATATCTTTTC 937
 Db 226 ValLysValPheTrpIleProGluSerGluAsnProAspAspAspLysIleTyrPhePhe 245
 QY 938 TTCCGTGAAATGCATAGATGGAGAACACTCT---GGAAAGCTACTCACCTCAGATTA 994
 Db 246 PheArgGluThrAlaValGluAlaAlaProAlaLeuGlyArgLeuSerValSerArgVal 265
 QY 995 GGTCAATATCAAGAAATGACTTTTGGAGGACACAGACTCTGCTCAATAAATGGACACA 1054
 Db 266 GlyGlnIleCysArgAsnAspValGlyGlnArgSerLeuValAsnLysTrpThrThr 285
 QY 1055 TTCTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAATGGCATGACACTATTTT 1114
 Db 286 PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe 304
 QY 1115 GATCACTGACAGGATGATTCCTAATGAATTTTAAAGATCTCTTAAATATCCAGTTGTATAT 1174
 Db 305 AspGlnLeuGlnAspValPheLeuLeuSerSerArgAspHisArgThrProLeuLeuTyr 324
 QY 1175 GGAGTGTTTACACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTGTATGATATGATG 1234
 Db 325 AlaValPheSerThrSerSerSerIlePheGlnGlySerAlaValCysValTyrSerMet 344
 QY 1235 AGTGATGTGAGAGGGTCTCTTGGTCCATATGCCACAGGATGGACCCCAACTATCAA 1294
 Db 345 AsnAspValArgAlaPheLeuGlyProPheAlaHisLysGluGlyProMetHisGln 364
 QY 1295 TGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACCTTGTCCGCAAAACA 1354
 Db 365 TrpValSerTyrGlnGlyArgValProTyrProArgProGlyMetCysProSerLysThr 384
 QY 1355 TTTGGTGGTTTGTACTCTACAAAGACCTTCTGATGATGTATTAACCTTTTCAGAGAT 1414
 Db 385 PheGlyThrPheSerSerThrLysAspPheProAspAspValIleGlnPheAlaArgAsn 404
 QY 1415 CATCCAGCATGTACAATCCAGTGTCTTCTATGAACAATGCCCAATAGTATCAAAACG 1474
 Db 405 HisProLeuMetTyrAsnSerValLeuProThrGlyGlyArgProLeuPheLeuVal 424
 QY 1475 GATGTAAATATCAATTTTACAAATTTGTCTAGACCGAGTGGATGACAGAGATGGACAG 1534
 Db 425 GlyAlaAsnTyrThrPheThrGlnIleAlaAlaAspArgValAlaAlaAspGlyHis 444
 QY 1535 TATGATGTATGTTTTCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCAATTCCT 1594
 Db 445 TyrAspValLeuPheIleGlyThrAspValGlyThrValLeuLysValIleSerValPro 464
 QY 1595 AAGGAGACTTCTGATGATTTTGAAGAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAA 1654
 Db 465 LysGlySerArgProSerAlaGluGlyLeuLeuLeuGluLeuHisValPheGluAsp 484
 QY 1655 CCGACTGCTATTTCAGCAATGGAGCTTTTCCACTAAGCAGCAACAACTATATATTGGTTCA 1714
 Db 485 SerAlaAlaValThrSerMetGlnIleSerSerLysArgHisGlnLeuTyrValAlaSer 504
 QY 1715 ACGGCTGGGTTGCCAGCTCCCTTTTACCGGTGTGATATTTACGGGAAGCGGTGCT 1774
 Db 505 ArgSerAlaAlaGlnIleAlaLeuHisArgCysAlaAlaHisGlyArgValCysThr 524

```

QY 1775 GAGTGTTCCTGCTCCGAGACCCCTTACTGCTGCTGGATGGTCTCTGCATGTTCTCGCTAT 1834
Db 525 GluCysLeuAlaArgAspProTyrCysAlaTrpAspGlyValAlaCysThrArgPhe 544
QY 1835 TTTCCACCTGCAAGAGAGCGCACAGACGACAGATATAGAAATGGAGACCCACTGACT 1894
Db 545 GlnProSerAlaLysArgPheArgGlnAspValArgAsnGlyAspProSerThr 564
QY 1895 CACTGTTACAGATTACACCATGATATACCATGCGCACGCCCTGAA----- 1942
Db 565 LeuCysSer-----GlyAspSerSerArgProAlaLeuLeu 576
QY 1943 GAGAGATCATCTATGCTGTAGAGATAGTAGACATATTTTGGATGCGAGTCGAGTCG 2002
Db 577 GluHisLysValPheGlyValGluGlySerSerAlaPheLeuGluCysGluProArgSer 596
QY 2003 CAGAGAGCGCTGCTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGAGATC 2062
Db 597 LeuGlnAlaArgValGluTrpThrPheGlnArgAlaGlyValThrAlaHisThrGlnVal 616
QY 2063 AGAGTGTGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAACAG 2122
Db 617 LeuAlaGluGluArgThrGluArgThrAlaArgGlyLeuLeuLeuArgArgLeuArg 636
QY 2123 AAGGATTCAGGCATTTACCTCTGCCATGCGGTGGACATGGCTTATACAAACTTCTT 2182
Db 637 ArgAspSerGlyValTyrLeuCysAlaValGluGlnGlyPheThrGlnProLeuArg 656
QY 2183 AAGGTAAACCTGGAAGTCAATGACACAGACAGCATTTTGGAGAACTTCTTCATAAAGATGAT 2242
Db 657 ArgLeuSerLeuHisValLeuSerAlaThrGlnAlaGluArgLeu----- 671
QY 2243 GATGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAGTTC 2302
Db 672 -----AlaArgAlaGluGluAlaAlaProAlaAlaProGlyProGlySerLeu 687
QY 2303 TGTACAGACTTCTAGCTCATCACCACCCCACTCTCAACAGTGTGATGATGTC 2362
Db 688 TrpTyrArgAspPheLeuGlnLeuVal----- 696
QY 2363 TGTGAACAAGTTTGGAAAGGGAGCGCAACCAACGTCGCAAGGCGGACATACCCCA 2422
Db 697 -----GluProGly-----GlyGly 701
QY 2423 GGAACAGTAAACAAATGGAAG-----CACTTA 2449
Db 702 GlySerAlaAsnSerLeuArgMetCysArgProGlnProAlaLeuGlnSerLeuProLeu 721
QY 2450 CAAGAAATTAAGAAAGGTAGAACAGGAGGAGGCCAC-----GAATTTCGAGGG 2497
Db 722 GluSerArgArgLysGlyArgAsnArgThrHisAlaProGluProArgAlaGluArg 741
QY 2498 GCACCCAGAGT 2509
Db 742 GlyProArgSer 745

```

RESULT 11

```

SM3B_MOUSE
ID SM3B_MOUSE STANDARD; PRT; 748 AA.
AC Q6217;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin A) (Sema A).
GN SEMA3B OR SEMA OR SEMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;

```

```

RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948 (1995).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; X85990; CAA59982.1; -.
DR PIR; I48744; I48744.
DR MGD; MGI:107561; Sema3b.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 748 SEMAPHORIN 3B.
FT DOMAIN 55 496 SEMA.
FT DOMAIN 561 659 IG-LIKE C2-TYPE.
FT DOMAIN 698 702 POLY-GLY.
FT DOMAIN 723 743 ARG-RICH (BASIC).
FT DISULFID 643 709 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 748 AA; 82894 MW; 1866B7D2397C9305 CRC64;

```

```

Alignment Scores:
Pred. No.: 2.88e-156 Length: 748
Score: 2059.50 Matches: 400
Percent Similarity: 66.93% Conservative: 112
Best Local Similarity: 52.29% Mismatches: 184
Query Match: 42.06% Indels: 69
DB: 1 Gaps: 15

```

US-09-774-490-1 (1-2709) x SM3B_MOUSE (1-748)

```

QY 281 AATGTGCCAAGCTGAATATCTTACAAAGAAATGTTGGAATCCAAACATGTGATCACT 340
Db 27 AsnLeuProArgLeuArgLeuSerPheGlnGluLeuGlnAlaArgHisGlyValArgThr 46
QY 341 TTCAATGGCTTGGCCCAACAGTCCAGTTATCATCTTCTTCTTGGATGAGAACGAGT 400
Db 47 PheArg---LeuGluArgThrCysCysTyrGluAlaLeuLeuValAspGluGluArgGly 65
QY 401 AGGTGTATGTGGAGCAAGATCACATATTTTCATTCGACCTGGTTAATATC---AAG 457
Db 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuSerLeuAsnIleSerLys 85
QY 458 GATTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCT 517

```

```
Db 86 ArgAlaLysLeuValAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
QY 518 GGAAAGACATCTGAAAGAAATGTGCTAATTTTCATCAAGTACTTAAGGCATATAATCAG 577
Db 106 GlyLysAspIleGlyThrGluCysMetAsnPheValArgLeuLeuHisAlaLysHis 125
QY 578 ACTCACTGTAGCCTGTGGAAAGCGGGGCTTTTCATCCAAATTCG 622
Db 126 ThrHisLeuLeuAlaCysArgThrGlyAlaPheHisProThrCysAlaLeuTrpArgTrp 145
QY 623 -----ACCTACATTGAAATTGGACATCATCTCGAGCACAATATTTTAAG 667
Db 146 AlaThrAlaGlyCysHisAlaSerThrGly-----ProGlu----- 158
QY 668 CTGGAGAACTCATATTTTGAAGACCGCGTGGGAAGAGTCCATATGACCTAAGCTGCTG 727
Db 159 -----LysLeuGluAspGlyLysGlyLysThrProTyrAspProArgHisArg 174
QY 728 ACAGCATCCCTTTTAATAGATGGAGAAATTAATCTCTGGAACTGCGAGCTGATTTATGGG 787
Db 175 ProProSerValLeuValGlyGluLeuTyrSerGlyValThrAlaAspLeuMetGly 194
QY 788 CGAGACTTGTCTATCTCCGAATCTCTGGCCACCACCCCAATCAGACAGCAGCAT 847
Db 195 ArgAspPheThrIlePheArgSerLeuGlyGlnAsnProSerLeuArgThrGluProHis 214
QY 848 GATTCAGGTGCTCAATGATCCAAAGTTTCATAGTGGCCACCTCATCTCAGAGAGTGAC 907
Db 215 AspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSerGlu 234
QY 908 AATCTGAAGATGACAAATGATATCTTTTCTCCGTGAAATGCAATAGATGAGGAACAC 967
Db 235 AsnProAspAspLysIleTyrPhePheArgGluSerAlaValGluAlaPro 254
QY 968 TCT---GGAAAGCTACTCACCTAGATAGTATGATGATGCAAGATGACTTTGAGGG 1024
Db 255 AlaMetGlyArgMetSerValSerArgValGlyGlnIleCysArgAsnAspLeuGly 274
QY 1025 CACGAAGTCTGGTGAATAAATGGCAACATTCCTCAAGCTCGCTGTGATTTGCTCAGTG 1084
Db 275 GlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerVal 294
QY 1085 CCAGTCCAAATGGCACTGACATCTTTGATGAACACGAGATGTATCTCAATGAAC 1144
Db 295 ProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeuSer 313
QY 1145 TTTAAGATCCTAAAATCCAGTTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATG 1204
Db 314 SerArgAspArgGlnThrProLeuLeuTyrAlaValPheSerThrSerSerGlyValPhe 333
QY 1205 AAGGGATCAGCCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1264
Db 334 GlnGlySerAlaValCysValLysSerMetAsnAspValArgAlaPheLeuGlyPro 353
QY 1265 TATGCCCAAGGATGGACCCCACTCAATCGGTCGCTTATCAAGGAGAGTCCCTTAT 1324
Db 354 LeuProHisLysGluGlyProThrHisGlnTrpValSerTyrGlnGlyArgValProTyr 373
QY 1325 CCACGGCAGGAACCTGTCCGAGCAAAACATTTGTTGTTTGTGACTCTCAAGAGACCTT 1384
Db 374 ProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAspPhe 393
QY 1385 CCTGATGATGTTAATACCTTTTCAAGAGTCTATCCAGCCATGTACATCCAGTGTTCCT 1444
Db 394 ProAspAspValIleGlnPheGlyArgAsnHisProLeuMetCysAsnProValLeuPro 413
QY 1445 ATGAACAATCGCCCAATAGTATCAAAACCGGATGTAATTTATCAATAATTTACACAAATGTC 1504
Db 414 MetGlyGlyArgProLeuPheLeuGlnValGlyAlaGlyTyrThrPheThrGlnIleAla 433
QY 1505 GTAGACCGAGTGGATGAGCAAGATGAGATGATGATGATGATGATGATGATGATGATGATG 1564
```

```
Db 434 AlaAspArgValAlaAlaAlaAspGlyHisTyrAspValLeuPheIleGlyThrAspVal 453
QY 1565 GGGACCGTCTTAAAGTAGTATTTCAATCTTAAGGAGACTTGTATGATTATGAAGAGGTT 1624
Db 454 GlyThrValLeuLysValIleSerValProLysGlyArgArgProAsnSerGluGlyLeu 473
QY 1625 CTGCTGGAAAGAAATGACAGATTTTTCGGGAACCGACTGCTATTTCAGCAATGAGCTTCC 1684
Db 474 LeuLeuGluGluLeuGlnValPheGluAspSerAlaAlaIleThrSerMetGlnIleSer 493
QY 1685 ACTAAGCAGCAACACTATATATGTTCAACGGCTGGGTGGCCAGCTCCCTTTTACAC 1744
Db 494 SerLysArgGlnGlnLeuTyrValAlaSerArgAlaAlaValAlaGlnIleAlaLeuHis 513
QY 1745 CGGTGTGATATTTACGGAAAGCGTGTCTGAGTGTCTGCTCGCCGAGACCTTACTGT 1804
Db 514 ArgCysThrAlaLeuGlyArgAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 533
QY 1805 GCTTGGGATGTTCTGCAATGTTCTCGTATTTTCTCCACTGCAAGAGACGACACAGCA 1864
Db 534 AlaTrpAspGlySerAlaCysThrArgPheGlnProThrAlaLysArgArgPheArg 553
QY 1865 CAAGATATAAGAAATGGAGACCTGACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1924
Db 554 GlnAspIleArgAsnGlyAspProSerThrLeuCysSerGly-----AspSerSer 570
QY 1925 CATGCCACAGCCCTGAAGAGAGATCATCTATGTTGTAGAGAAATAGTAGCATTCTTTG 1984
Db 571 HisSerValLeuLeuGluLysLysValLeu---GlyValGluSerGlySerAlaPheLeu 589
QY 1985 GAATGCACTCGAAGTCCGAGAGCGCTGTGTTTGGCAATTCAGAGCGGCAATGAA 2044
Db 590 GluCysGluProArgSerLeuGlnAlaHisValGlnTrpThrPheGlnGlyAlaGlyGlu 609
QY 2045 GAGCGAAAGAGAGATCAGAGTGCATCATATCATCAGCAGCAGATCAGGCCCTCTG 2104
Db 610 AlaAlaHisThrGlnValLeuAlaGluArgValGluArgThrAlaArgGlyLeuLeu 629
QY 2105 CTACGTAGCTTACAAACAGAGGATTCAGGCAATTTACCTCTGCTGCTGCTGCTGCTGCTG 2164
Db 630 LeuArgGlyLeuArgArgGlnAspSerGlyValTyrLeuCysValAlaValGluGlnGly 649
QY 2165 TTCATACAACTCTTTTAAAGTAACCTGGAAGTCACTTGCACAGAGCATTTTGAAGAA 2224
Db 650 PheSerGlnProLeuArgArgLeuValLeuHisValLeuSerAlaAlaGlnAlaGluArg 669
QY 2225 CTTCTTCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2284
Db 670 Leu-----AlaArgAlaGluGluAlaAlaProAla 680
QY 2285 ACACCTAGCCAGAGGCTCTGTTACAGAGCTTTCATGACAGCTCATC-----AACCAC 2335
Db 681 ProProGlyProLysLeuTyrTyrArgAspPheLeuGlnLeuValGluProGlyGlyGly 700
QY 2336 CCCAATCTCAACAGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2395
Db 701 GlyGlyAlaAsnSerLeu-----ArgMet 708
QY 2396 CGTGGCAAGAGCCAGGACATACCCAGGGACAGTAACAATGGAAGCAC----- 2446
Db 709 CysArgProGlnProGlyHis-----HisSerValAla 719
QY 2447 TTACAAGAAATAAGAAAGTAGAAACAGGAGGACCCAC-----GAATTTGAG 2494
Db 720 AlaAspSerArgArgLysGlyArgAsnArgMetHisValSerGluLeuArgAlaGlu 739
QY 2495 AGGCACCCAGGAGT 2509
Db 740 ArgGlyProArgSer 744
```

RESULT 12
SM3E_HUMAN PRT; 775 AA.
ID SM3E_HUMAN STANDARD;

KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 775 SEMAPHORIN 3E.

FT DOMAIN 241 540 SEMA.

FT DOMAIN 581 669 IG-LIKE C2-TYPE.

FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).

FT DISULFID 654 729 BY SIMILARITY.

FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 775 AA; 89503 MW; BA1690093D28F40 CRC64;

Alignment Scores:

Pred. No.: 4,12e-153 Length: 775

Score: 2020.00 Matches: 367

Percent Similarity: 67.62% Conservative: 155

Best Local Similarity: 47.54% Mismatches: 222

Query Match: 41.26% Indels: 18

DB: 1 Gaps: 8

US-09-774-490-1 (1-2709) x SM3E_MOUSE (1-775)

QY 218 ATTGCTGTCTTCTGGGAGTATTACTTACGACGAGCAAACTATCAGATGGGAG 277

DB 8 LeuThrLeuLeuLeuProGlyHisLeuLeuGluLeuThrProGlyHisSerAlaAsn 27

QY 278 AACAAATGTCAGGCTGAATATCTCAAAAGAAATGTTGGAATCCAACTATGATC 337

DB 28 ProSerTyrProArgLeuArgLeuSerHisLysGluLeuLeuGluLeuAsnArgThrSer 47

QY 338 ACTTCAATGGCTGGCCAAAGCTCCAGTTATCATCTTACCTTCTTTGGATGAGAACGG 397

DB 48 IlePheGlnSerProLeuGlyPheLeuAspLeuLeuHisThrMetLeuLeuAspGluTyrGln 67

QY 398 AGTAGCTGTATGTCGACCAAGATCATATTTTCACTTCGACTGCTTAATATCAAG 457

DB 68 GluArgLeuPheValGlyGlyArgLeuValTyrSerLeuAsnGluArgValSer 87

QY 458 GAT---TTTCAAAAGATTGTGGCCAGTATCTTACACGACGAGATGAATGCAAGTGG 514

DB 88 AspGlyTyrArgGluIleTyrTrpProSerThrAlaValLysValGluCysIleMet 107

QY 515 GCTGGAAGACATCTCGAAGATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAT 574

DB 108 LysGlyLysAsp---AlaAsnGluCysAlaAsnTyrIleArgValLeuHisHisTyrAsn 126

QY 575 CAGACTCACTGTACGCTGTGGACGGGGCTTTTCATCCCAATTTGCACCTACATTGAA 634

DB 127 ArgThrHisLeuLeuThrCysAlaThrGlyAlaPheAspProHisCysAlaPheIleArg 146

QY 635 ATTGACATCATCTCGAGCAATATTTTAAAGCTGGAGAACTCAATTTTCAAAACGGC 694

DB 147 ValGlyHisHisSerGluGluProLeuPheHisLeuGluSerHisArgSerGluArgGly 166

QY 695 CQTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAA 754

DB 167 ArgGlyArgCysProPheAspProAsnSerSerPheValSerThrLeuValGlyAsnGlu 186

QY 755 TTATCTCTGGACTGCACCTGATTTTATGGGCGGAGCTTTGCTATCTTCCGACTCTT 814

DB 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpGlyArgAspSerAlaIlePheArgSerMet 206

QY 815 GGGCACCAACCCCAATCAGGACAGACGACGATGATTTCCAGGTGGCTCAATGATCCAAAG 874

DB 207 GlyLysLeuGlyHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 226

QY 875 TTCATTAGTCCCACTCATCTCAGAGAGTGCATTCCTTGAAGATGACCAAACTACTTT 934

DB 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysMetTyrPhe 246

QY 935 TTCTTCGTGAAATGCAATAGATGAGACACACTCTGGAAGAGCTACTCAGCCTAGAATA 994

DB 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAsnAlaHisThrIleTyrThrArgVal 266

QY 995 GGTGAGATATCAAGAATGACTTTTGGAGGGCACAGAAGTCTGGTGAATAAATAATGGAACA 1054

DB 267 GlyArgLeuCysValAsnAspMetGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286

QY 1055 TTCCTCAAGCTCCTGATTTGCTCAGTCCAGGTCCAATGCGCATTTGACATCATTTT 1114

DB 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306

QY 1115 GATCAACTGACGAGTATCTCTAATGAATTTTAAAGATCTTAAATAATCCAGTTCTATAT 1174

DB 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspProLysAsnProValIlePhe 326

QY 1175 GGAGTGTTCACACTTCCAGTAACATTTTCAAGGATCAGCCGTGTGTATGATATGATCATG 1234

DB 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaValCysValTyrHisMet 346

QY 1235 AGTGATGTGACAGGGTGTCTCTTGGTCCATATGCCACAGGGATGGACCCCACTATCAA 1294

DB 347 SerSerIleArgGluAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366

QY 1295 TGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGCAACTTGTCCAGCAAAAACA 1354

DB 367 TrpSerLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386

QY 1355 TTTGGTGTG---TTTGACTCTCAAAAGACCTTCTCTGATGATGTTTAACTTTTCCAGA 1411

DB 387 AsnGlyGlyLysTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406

QY 1412 AGTCATCCAGCCATGTCACATCCAGTGTTCCTTATGAACAATCGCCCAATAGTATCAA 1471

DB 407 IleAspProLeuMetTyrGlnProIleLysProValHisLysLysProIleLeuValLys 426

QY 1472 ACGATGTAAATTTATCAATTTTACAAATTTCTGTAGACCGAGTGGATGCAAGATGGA 1531

DB 427 ThrAspGlyLysTyrAsnLeuArgGlnLeuAlaValAspArgValGluAlaGluAspGly 446

QY 1532 CAGTATGATGTTATGTTTATCGGAACAGATTTGGGACCGTCTTCTTAAAGTATGTTCAAT 1591

DB 447 GlnTyrAspValLeuPheIleGlyThrAspThrGlyIleValLeuLysValIleThrIle 466

QY 1592 CTAAGGAGACTTGGTATGATTTTAGACAGCTTCTGCTGGAGAAATCACAGTTTTCGG 1651

DB 467 TyrAsnGlnGluThrGluTrpMetGluGluValIleLeuGluGluGlnIlePheLys 486

QY 1652 GAACCGACTGTATTTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATGTT 1711

DB 487 AspProAlaProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506

QY 1712 TCAACGGCTGGGTGGCCAGCTCCCTTTACCGGTGTGATATTTACGGGAAGCGGTGT 1771

DB 507 SerAlaSerAlaValAlaGlnValArgPheHisCysAspMetTyrGlySerAlaCys 526

QY 1772 CQTGAGTGTTCCTCGCCGAGACCTTACTGCTGGTGGATGTTCTGCAATGTTCTCGC 1831

DB 527 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546

QY 1832 TATTTTCCACT-----GCAAGAGACGCAACAGCAACAGATATTAAGAAATGGA 1882

DB 547 TyrTyrProThrGlyAlaHisAlaLysArgArgPheArgAGlnAspValArgHisGly 566

QY 1883 GACCCACTGACTCTGTTTCAGACTTACCATGATATACCATGGCCACAGCCCTGAA 1942

DB 567 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspArgThrGlu 586

QY 1943 GAGAGATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGATGTCAGTCCGAGTCCG 2002

DB 587 GluArgLeuAlaTyrGlyIleGluSerAsnSerThrLeuLeuGluCysThrProArgSer 606

QY 2003 CAGAGAGCGCTGGTCTATTGGCAATTTCCAGAGGCGAAATGAAGAGCGAAAAAGAGATC 2062

607	LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgAspValArgLysGluGluVal	626
2063	AGAGTGATGATCATATCATCATGACGACAGATCAAGGCTTCTGCTACGTAGTCTCAACACG	2122
627	LysThrAspAspArgValValLysMetAspLeuglyLeuLeuPheLeuArgValArgLys	646
2123	AAGGATTTCAGCAATTACCTCTGCATCGCGTGGAAACATGGTGGTTCATACAACTTCCTT	2182
647	SerAspAlaGlyThrTyrrPheCysGlnThrValGluHisAsnPheValHisThrValArg	666
2183	AAGGTAACCCCTGGAGTCATTGTACACAGACGATTTGGAAAGAACTCTTCCTATAAGATCAT	2242
667	LysIleThrLeuGluValValGluGluHisLysValGluGlyMetPheHisLysAspHis	686
2243	GATGGAGATGGCTCTAAGACCAAGAATAATGTCC-----AATGATCATGACACCT	2290
687	GluGluGlu-----ArgHisLysMetProCysProLeuSerGlyMetSerGln	704
2291	AGCCAGAAGGTCTGGTACAGAGACTTCATGCAGCTCATCAACNCCCATCTCAACACG	2350
705	GlyThrLysProTrpTrpLysGluPheLeuGlnLeuIleGlyTrpSerAsnPheGlnArg	724
2351	ATGATCAGATTCTGTGAACAAGTTTGAAAGGACCGCAAACTCTCGCAAGGCCA	2410
725	ValGluGluTrpCysGluLysValTrpCysThrAspLysArgLysLysLeuLysMet	744
2411	GGACATACCCCGGGAACAGTAAACAATGAAGCAGCTTACCAAGAAATAAGAAAGTGA	2470
745	-----SerProSerLysTrpLysTyrAlaAsnProGlnGluLys--Arg	758
2471	AACAGGAGGACCCACGAATTTGAGAGGCCACCCAGG	2506
759	LeuArgSerLysAlaGluHisPheArgLeuProArg	770

RESULT 14

SM3E_CHICK					
ID	SM3E_CHICK	STANDARD;	PRT;	785	AA.
AC	Q42237;	Q90666;			
DT	30-MAY-2000	(Rel. 39., Created)			
DT	30-MAY-2000	(Rel. 39., Last sequence update)			
DT	18-OCT-2001	(Rel. 40., Last annotation update)			
DE	Semaphorin 3E precursor (Collapsin-5) (COLL-5).				
GN	SM3A3E OR COLL5.				
QS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCBI_TaxID=9031;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fetal brain;				
RX	MEDLINE=97470887; PubMed=9331347;				
RA	Feiner L., Koppel A.M., Kobayashi H., Raper J.A.;				
RA	"Secreted chick semaphorins bind recombinant neuropilin with similar				
RT	affinities but bind different subsets of neurons in situ.";				
RL	Neuron 19:539-545(1997).				
RN	[2]				
RN	SEQUENCE OF 244-543 FROM N.A.				
RP	MEDLINE=95329269; PubMed=7605628;				
RA	Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;				
RT	"A family of molecules related to collapsin in the embryonic chick				
RT	nervous system.";				
RL	Neuron 14:1131-1140(1995).				
CC	-1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH				
CC	CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC				
CC	NEURONAL POPULATIONS. BINDS TO NEUROFILIN.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- TISSUE SPECIFICITY: COLLAPSIN-1, -2, -3, AND -5 BIND TO				
CC	OVERLAPPING BUT DISTINCT AXON TRACTS.				
CC	-1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY				
CC	THIRD OF THE PROTEIN.				
CC	-1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.				

CC	-I- SIMILARITY: Contains 1 Sema domain.
CC	-I- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AF022947; AAB80952.1; --
DR	EMBL; U28243; AAA85899.1; --
DR	InterPro; IPR003659; Flexin-like.
DR	InterPro; IPR001627; Sema.
DR	Pfam; PF01403; Sema; 1.
DR	SMART; SM00423; PSI; 1.
DR	SMART; SM00630; Sema; 1.
KW	Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW	developmental protein; Glycoprotein.
FT	SIGNAL 1 25 POTENTIAL.
FT	CHAIN 1 25 SEMAPHORIN 3E.
FT	DOMAIN 26 785 Sema.
FT	DOMAIN 244 544 IG-LIKE C2-TYPE.
FT	DOMAIN 651 740 ARG/LYS-RICH (BASIC).
ET	DOMAIN 741 780 BY SIMILARITY.
FT	DISULFID 658 733 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 600 600 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT 246 246 N -> D (IN REF. 2).
FT	CONFLICT 248 248 V -> I (IN REF. 2).
FT	CONFLICT 250 250 L -> F (IN REF. 2).
SQ	SEQUENCE 785 AA; 90978 MW; B551EBF717630632 CRC64;
Alignment Scores:	
Pred. No.:	5,9e-151 Length: 785
Score:	1993.00 Matches: 367
Percent Similarity:	67.26% Conservative: 161
Best Local Similarity:	46.75% Mismatches: 223
Query Match:	40.71% Indels: 34
DB:	1 Gaps: 10
US-09-774-490-1 (1-2709) x SM3E_CHICK (1-785)	
Qy	218 ATTTGTCGTCTTTCTGGCGAGTATTACTACAGCAAGACAACTATCAGAAATGGGAAG 277
Db	12 LeuileuAlaLeuCysGlyLeuLeuLeuLeuLeuProAlaGlyTyrHisAlaThrAsp 31
Qy	278 AACAAATGTGCAGGCTGAAATTTATCTCTCAAGAATAATGTGGAAATCCAAATGTGATC 337
Db	32 ThrArgGlnProArgLeuArgLeuSerHisLysGluLeuTrpAspLeuAsnArgThrSer 51
Qy	338 ACTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTCTCTTTTGGATGGGAACGG 397
Db	52 ValPheHisSerProPheGlyTyrLeuGlyLeuHisIleMetLeuLeuAspGluTyrGln 71
Qy	398 AGTAGGCTGTATCTTGCAAGAGATCATATTTTCATTTTCGACCTGGTTAATATC--- 454
Db	72 GluArgLeuPheValGlyGlyArgAspLeuLeuTyrSerLeuSerLeuAspArgIleSer 91
Qy	455 AAGGATTTTCAAAGATTGTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514
Db	92 AsnAsnTyrArgGluIleHisTrpProSerThrProLeuGlnAlaGluGlyCysIleIle 111
Qy	515 GCTGGAAAGACATCTCTGAAAGAAATGTGCTAATTTTCATCAAGGTACTTAAAGGCATATAAT 574
Db	112 LysGlyArgAsp---AlaAspGluCysAlaAsnTyrValArgValLeuHisArgTyrAsn 130
Qy	575 CAGACTCACTTGTACGCTGTGGACGGGGCTTTTCATCCAAATTTGCACCTACATTGAA 634
Db	131 ArgThrHisLeuLeuAlaCysGlyThrGlyAlaPheAspProValCysThrPheIleArg 150
Qy	635 ATTGGACATCATCTCTGAGGACAAATATTTTTTAAGCTGGAGAACTCACATTTTGAACACGGC 694

151	Db	ValGlyHisProSerGluAspHisLeuPheGlnLeuGluSerHisLysPheGluArgGly	170
695	Qy	CGTGGGAAGAGTCCATATGACCTTAAGCTGTCGACAGCATCCCTTTTAAATAGATGGAGAA	754
171	Db	ArgGlyArgCysProPheAspProThrSerSerPheThrSerIleLeuIleGlyGlu	190
755	Qy	TTATACTCTGGAACTCGACGTGATTTTATGGCGGAGACTTTGCTACTCTCCGAACCTCTT	814
191	Db	LeuPheThrGlyLeuTyrSerAspTyrTrpGlyArgAspAlaValPheArgThrMet	210
815	Qy	GGGCACCACCAACCAATCAGACAGAGAGCATGATTCCAGGTGGCTCAATGATGCCAAG	874
211	Db	AsnArgMetAlaHisLeuLeuArgThrGluProAspSerGluHisLeuLeuLysGluProLys	230
875	Qy	TTCAATTAGTCCCACTCATCTCAGAGAGTCACAACTCCTGAGATGACAAAGTACTACTTT	934
231	Db	PheValGlySerTyrMetIleProAspAsnGluAspHisAspAspAsnLysValTyrLeu	250
935	Qy	TTCTTCCTCGTGAATAAGCAATAGATGGAGAACACTCTCGAAAGCTACTCAGCTAGAAATA	994
251	Db	PhePheThrGluLysAlaLeuGluAlaGluThrSerThrHisAlaIleTyrThrArgVal	270
995	Qy	GGTCAGATATGCAAGAATGACTTTGGAGGGCAGACAAAGTCTGGTGAATAAATGGACAACA	1054
271	Db	GlyArgValCysValAsnAspMetGlyGlyGlnArgIleValValAsnLysTyrSerThr	290
1055	Qy	TTCTCTAAAGCTCGTCTGATTGCTTCAGTGCAGGTCCAAATGGCATTGACACTCATTTT	1114
291	Db	PheLeuLysThrArgLeuValCysSerValProGlyArgAsnGlyIleAspThrHisPhe	310
1115	Qy	GATGAACCTGCAGGATGATTCTCTAATCAACTTTTAAAGATCTTAAAGATCCAGTTGTATAT	1174
311	Db	AspGluLeuGluAspValPheLeuLeuGlnThrArgAspAsnLysAsnProValIlePhe	330
1175	Qy	GGAGTGTTACGACTCCAGTAAACATTTTCAAGGGATCAGCCGGTGTATGTATPAGCATG	1234
331	Db	GlyLeuPheSerThrThrSerAsnIlePheArgGlyTyrAlaIleCysValTyrHisMet	350
1235	Qy	AGTCATGTGAGAAGGGTTCCTTGGTTCATATGCCACAGGGATGAGCAACCACTATCAA	1294
351	Db	AlaIleValArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrTyr	370
1295	Qy	TGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGCGGCAGAACTCTGCCAGCAAAACA	1354
371	Db	TrpAlaLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal	390
1355	Qy	TTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTCGATGATGTTATTAACCTTTGCAAGA	1411
391	Db	AsnGlyGlyLeuTyrThrThrThrLysAspTyrProAspGluAlaValHisPheAlaArg	410
1412	Qy	AGTCATCCAGCCATGTACATCCAGTGTTCCTTATGAACAACTCCGCCAATAGTGTCAAA	1471
411	Db	SerHisProLeuMetTyrGlnProIleLysProValHisLysArgProIleLeuValLys	430
1472	Qy	ACGGATGTAAATTATCAATTATACAAATTCCTAGACCGAGTGGAGTGCAGAAAGATGGA	1531
431	Db	ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly	450
1532	Qy	CAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTCTTAAAGTAGTTCAAAT	1591
451	Db	GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle	470
1592	Qy	CCTAAGAGACTTGGTATGATTTAGAGAGGTTCCTGCTGGAGAAATGACAGTTTTCGG	1651
471	Db	TyrAsnGlnGlnThrGluSerMetGluGluValIleLeuGluGluLeuValPheLys	490
1652	Qy	GAACCGACTGCTATTTCCAGCAATCGAGCTTTCCACTAAGCAGCAACAATATATATGGT	1711
491	Db	ValProIleProIleLeuSerMetCulIleSerSerLysArgGlnGlnLeuTyrIleGly	510
1712	Qy	TCAACGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTTACGGGAAGCGTGT	1771

Db 511 ThrGluSerValIleAlaGlnValLysPheHisGlnCysAspMetTyrGlyThrAlaCys 530
 Qy 1772 GCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTTGGGATCGTTCGATGTTCTCGC 1831
 Db 531 AlaAspCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 550
 Qy 1832 TATTPTTCCCCT-----GCAAGAGAGACGCAACAGACGCAAGATATAAGAAATGGA 1882
 Db 551 TyrTyrProThrGlyMetGlnAlaLysArgArgPheArgArgGlnAspValArgHisGly 570
 Qy 1883 GACCCACTGACTCAGTTCCTGAGCTTACACCATGATTAATCACCATGCCACACCCCTGAA 1942
 Db 571 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheIleGlyGluValLeuGluLysThrGlu 590
 Qy 1943 GAGAGATCATCTATCGTGTAGAGAATAGTAGACACATTTTGGAAATGCAGTCGCGAAGTCG 2002
 Db 591 GluArgLeuValTyrGlyIleGluTyrAsnSerThrLeuLeuGluTyrThrProArgThr 610
 Qy 2003 CAGAGAGCGTGTCTATTGGCAATTCAGAGCGGAAATCAAGAGCGGAAAGAGAGATG 2062
 Db 611 LeuGlnAlaLysValAsnTrpPheValGlnArgAlaHisGluThrLysLysGluGluVal 630
 Qy 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTAGCTAGTCTACACAG 2122
 Db 631 LysThrAspGluArgIleIleLysMetAspLeuGlyLeuLeuPheLeuLysLeuHisArg 650
 Qy 2123 AAGGATTCAGGCAATTCCTCTGCCATCGGTGGAAATCGGTTTCATACAACTCTTCTT 2182
 Db 651 LeuAspAlaGlyThrTyrPheCysGlnThrValGluHisSerIleValHisThrValArg 670
 Qy 2183 AAGGTAACCCCTGGAAGTCATTGTACACAGAGCATTTTGGAAAGAACTTCTTCATAAAGATGAT 2242
 Db 671 LysIleThrLeuGluIleValGluGluGluArgValAspGluMetPheSerLysAspTyr 690
 Qy 2243 GATGGAGATGGCTCTAAGACCAAGAANAATGTCGAATAGCATG---ACACCTACCCAG--- 2296
 Db 691 Glu-----GluGluIleSerHisLysMetProCysProMetGlnSer 704
 Qy 2297 -----AAGGTCCTGTCACAGAGACTTCATGCAGCTCATCAACACACCC 2338
 Db 705 AsnIleProGlnValSerLysProTyrTyrLysGluPheLeuGlnLeuIleGlyTyrSer 724
 Qy 2339 AATCTCAACACGATGATGAGTTCCTGTGAACAAGTTTGGAAAGGACCGCAACCAACG 2398
 Db 725 AsnPheGlnArgValGluGluTyrCysGluLysValTrpCysThrAspLysLysArgLys 744
 Qy 2399 CGGAAAGGCCAGGACATACCCAGGGAACAGTAACAATGGAGCACTTA----- 2449
 Db 745 LysLeuLysMet-----SerProSerLysTrpLysTyrAlaAsnProGln 759
 Qy 2450 -----CAAGAAATAAGAAGGTAGAAACAGGAGGACCCACGAAATTGAGAGG 2497
 Db 760 GluLysArgGlnAspGlnGluLysLysAlaArgIleArgProGluHis-----TyrArg 777
 Qy 2498 GCACCCAGGAGTGC 2512
 Db 778 LeuProArgAsnIle 782
 RESULT 15
 SM3F HUMAN
 ID SM3F HUMAN STANDARD; PRT; 785 AA.
 AC Q13275; Q13274; Q13372; Q15704;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV) (Sema III/F).
 GN SEMA3F.
 OS Homo sapiens. (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Petal brain;
RX MEDLINE=96226360; PubMed=8649831;
RA Roche J., Boldog F., Robinson M., Robinson L., Varella-Garcia M.,
RA Swanton M., Waggoner B., Fishel R., Franklin W., Gemmill R.,
RA Drabkin H.;
RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
RT human semaphorin.";
RL Oncogene 12:1289-1297(1996).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MET-503.
RC TISSUE=Petal brain;
RX MEDLINE=96230324; PubMed=8786119;
RA Xiang R.-H., Hensel C.H., Garcia D.K., Carlson H.C., Kok K.,
RA Daly M.C., Kerbacher K., van den Berg A., Veldhuis P., Buys C.H.C.M.,
RA Naylor S.L.;
RT "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
RT 3p21, a region deleted in lung cancer.";
RL Genomics 32:39-48(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Nelson J., Biewald T.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 394-436 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh P.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
CC -!- FUNCTION: MAY PLAY A ROLE IN CELL MOTILITY AND CELL ADHESION.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
CC VARIETY OF NEURAL AND NONNEURAL TISSUES. THERE IS HIGH EXPRESSION
CC IN MAMMARY GLAND, KIDNEY, FETAL BRAIN, AND LUNG AND LOWER
CC EXPRESSION IN HEART AND LIVER.
CC -!- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 10.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33920; AAC50568.1; -;
DR EMBL; U38276; AAB18276.1; -;
DR EMBL; AC000063; BAB46344.1; -;
DR EMBL; U32171; AAB06011.1; -;
DR EMBL; U32172; AAB06012.1; -;
DR Genew; HGNC:10728; SEMA3F.
DR MIM; 601124; -;
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
KW Polymorphism. 1 18
FT SIGNAL 19 785
FT CHAIN

FT	DOMAIN	272	569	SEMA.
FT	DOMAIN	605	690	IG-LIKE C2-TYPE.
FT	DOMAIN	758	779	ARG/LYS-RICH (BASIC).
FT	DISULFID	678	746	BY SIMILARITY.
FT	CARBOHYD	53	53	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	474	474	A -> G (in dbSNP:1046955).
FT	VARIANT	503	503	/FTID=VAR_011820.
FT	VARIANT	503	503	L -> M.
FT	CONFLICT	153	183	/FTID=VAR_008855.
FT	CONFLICT	270	270	MISSING (IN REF. 2).
FT	CONFLICT	473	473	MISSING (IN REF. 2).
FT	CONFLICT	473	473	A -> S (IN REF. 2).
SQ	SEQUENCE	785 AA;	88381 MW;	FE3FC796EEC1608E CRC64;

Alignment Scores:
Pred. No.: 9,99e-147 Length: 785
Score: 1940.00 Matches: 372
Percent Similarity: 63.98% Conservative: 136
Best Local Similarity: 46.85% Mismatches: 226
Query Match: 39.62% Indels: 60
DB: 1 Gaps: 11

US-09-774-490-1 (1-2709) x SM3F_HUMAN (1-785)

QY	227	CTTTCTGGGAGTATTACTTTTAC	---GCAAGAGCAAACTAT	-----CAGAATGGGAG	277
DB	7	LeuLeuTrpAlaSerLeuLeuThrGlyAlaTrpProSerPheProThrGlnAspHisLeu	26		
QY	278	AACATGTGCCAGCTGAAATTTATCTACAAAGAAATGTTGGAATCCACAAATGTGATC	337		
DB	27	ProAlaThrProArgValArgLeuSerPheLysGluLeuLysAlaThrGlyThrAlaHis	46		
QY	338	ACTTTCATGCTTGGCCACAGCTCCAGTATCATACCTCTCTTTCGATGAGGACGG	397		
DB	47	PhePheAsnPheLeuLeuAsnThrAspTyrArgIleLeuLeuLysAspIleAspHis	66		
QY	398	AGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATCAAG	457		
DB	67	AspArgMetTyrValGlySerLysAspTyrValLeuSerLeuAspLeuHisAspIleAsn	86		
QY	458	GATTTTCAAGATGTG---TGCCAGTATCTTACACCAAGAGATGATGCAATGG	514		
DB	87	ArgGluProLeuIleIleHisTrpAlaAlaSerProGlnArgIleGluGluCysValLeu	106		
QY	515	GCTGAAAGACATCTCGAAGAGATGCTAATTTTCATCAAGTACTTAAGGCATATAAT	574		
DB	107	SerGlyLysAspValAsnGlyGluCysGlyAsnPheValArgLeuIleGlnProTrpAsn	126		
QY	575	CAGACTCACTGTACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATTCGA	634		
DB	127	ArgThrHisLeuTyrValCysGlyThrGlyAlaTyrAsnProMetCysThrTyrValAsn	146		
QY	635	ATTGACATCATCTCT	649		
DB	147	ArgGlyArgAlaGlnAlaThrProTrpThrGlnThrGlnAlaValArgGlyArgGly	166		
QY	650	-----GAGGACAAATATT	661		
DB	167	SerArgAlaThrAspGlyAlaLeuArgProMetProThrAlaProArgGlnAspTyrIle	186		
QY	662	TTTAAGCTGGAGAACTCACATTTTGAAGAGAGAGTCCATATGACCCCTAAG	721		
DB	187	PheTyrLeuGluProGluArgLeuGluSerGlyLysGlyCysProTyrAspProLys	206		
QY	722	CTGCTGACAGATCCCTTTTAAATAGATGGAGAAATATCTCTGGAATGACGTGATTTT	781		
DB	207	LeuAspThrAlaSerAlaLeuIleAsnGluLeuTyrAlaGlyValTyrIleAspPhe	226		
QY	782	ATGGGGCGAGACTTTGCTATCTCTCGAATCTTGGGACACACCCCAATCAGACAGAG	841		
DB	227	MetGlyThrAspAlaAlaIlePheArgThrLeuGlyLysGlnThrAlaMetArgThrAsp	246		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:36 ; Search time 28 Seconds
(without alignments)
9099.667 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatctttttttatcgtg.....aggctttttttcctaacc 2709

Scoring table: OLIGO Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Word size: 1

Total number of hits satisfying chosen parameters: 255550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09774490/runat_03082003.102927.4449/app_query.fasta_1.2887
-DB=Swissprot_41 -QFMT=fastan -SUFFIX=oligo.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490@cgn 1 1 33 @runat_03082003.102927.4449 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	1	SM3A_HUMAN
2	92	10.3	772	1	SM3A_MOUSE
3	92	10.3	772	1	SM3A_RAT
4	58	6.5	772	1	SM3A_CHICK
5	35	3.9	860	1	SM3A_BRARE
6	24	2.7	778	1	SM3B_BRARE
7	20	2.2	748	1	SM3B_MOUSE
8	16	1.8	764	1	SM22_BRARE
9	15	1.7	749	1	SM3B_HUMAN
10	15	1.7	751	1	SM3C_CHICK
11	15	1.7	751	1	SM3C_HUMAN
12	15	1.7	751	1	SM3C_MOUSE
13	14	1.6	777	1	SM3D_HUMAN
14	13	1.5	761	1	SM3D_CHICK
15	13	1.5	775	1	SM3E_HUMAN
16	13	1.5	775	1	SM3E_MOUSE
17	13	1.5	785	1	SM3E_CHICK
18	13	1.5	785	1	SM3F_HUMAN
1	771	86.1	771	1	Q14563 homo sapien
2	92	10.3	772	1	O08665 mus musculus
3	92	10.3	772	1	Q63548 rattus norv
4	58	6.5	772	1	Q90607 gallus gall
5	35	3.9	860	1	Q9w7j1 brachydanio
6	24	2.7	778	1	Q9w686 brachydanio
7	20	2.2	748	1	Q62177 mus musculus
8	16	1.8	764	1	Q9w696 brachydanio
9	15	1.7	749	1	Q13214 homo sapien
10	15	1.7	751	1	O42236 gallus gall
11	15	1.7	751	1	Q99985 homo sapien
12	15	1.7	751	1	Q62181 mus musculus
13	14	1.6	777	1	Q95025 mus musculus
14	13	1.5	761	1	Q90663 gallus gall
15	13	1.5	775	1	O15041 homo sapien
16	13	1.5	775	1	P70275 mus musculus
17	13	1.5	785	1	O42237 gallus gall
18	13	1.5	785	1	Q13275 homo sapien

08632	mus musculus	1	SM3F_MOUSE	785	1.5	13
Q64151	mus musculus	834	SM4C_MOUSE	1.1	10	
Q90665	gallus gall	295	SM4D_CHICK	1.0	21	
P09871	homo sapien	688	C1S_HUMAN	1.0	c	
Q62179	mus musculus	782	SM4B_MOUSE	1.0	23	
Q9npr2	homo sapien	832	SM4B_HUMAN	1.0	24	
O09126	mus musculus	861	SM4D_MOUSE	1.0	25	
Q92854	homo sapien	862	SM4D_HUMAN	1.0	26	
Q99613	homo sapien	913	IF38_HUMAN	1.0	27	
P32790	saccharomyc	1244	SLA1_YEAST	1.0	28	
Q15413	homo sapien	4870	RYR3_HUMAN	1.0	c	
Q92736	homo sapien	4967	RYR2_HUMAN	1.0	c	
P30957	oryctolagus	4969	RYR2_RABIT	1.0	31	
PL1902	escherichia	50	PND1_ECOLI	0.9	32	
P16477	escherichia	50	PND2_ECOLI	0.9	33	
P19466	bacillus su	75	MRB_BACSU	0.9	34	
O13062	trimeresur	257	VSPC_TRIGA	0.9	c	
P39570	bacillus su	368	GRBB_BACSU	0.9	c	
P50452	homo sapien	374	SPB8_HUMAN	0.9	37	
PL1614	canis faml	377	SH1D_CANFA	0.9	38	
P15628	soybean chl	462	IBMP_SOCMV	0.9	39	
P45161	h penicilli	479	PBP4_HAETN	0.9	c	
Q12627	klyveromyc	579	DLD1_KLULA	0.9	41	
O13773	schizosacch	697	YB9C_SCHPO	0.9	42	
Q62178	mus musculus	760	SM4A_MOUSE	0.9	43	
Q9h3s1	homo sapien	761	SM4A_HUMAN	0.9	44	
Q9whu7	mus musculus	837	SM4G_MOUSE	0.9	45	
Q9ntn9	homo sapien	838	SM4G_HUMAN	0.9	46	
P02574	petromyzon	966	FTB1_PETWA	0.9	47	
Q13591	homo sapien	1074	SM5A_HUMAN	0.9	48	
Q62217	mus musculus	1077	SM5A_MOUSE	0.9	49	
PL19176	pseudomonas	1409	RPOC_PSEPU	0.9	50	
PL1960	sus scrofa	5035	RYR1_PIG	0.9	c	
PL1716	oryctolagus	5037	RYR1_RABIT	0.9	c	
P21817	homo sapien	5038	RYR1_HUMAN	0.9	c	
Q82094	euglena gra	37	PSAJ_EUGGR	0.8	54	
Q82094	pyrobaculum	54	RS14_PYRAE	0.8	55	
P32040	synetococ	63	YPH2_SYNPF	0.8	c	
P33988	azorhizobiu	68	YRP3_AZOCA	0.8	57	
O28167	archaeoglob	99	YL13_ARCFU	0.8	58	
Q9pr15	columba liv	100	KRP1_COLLI	0.8	59	
Q93499	columba liv	100	KRP2_COLLI	0.8	60	
Q93500	columba liv	100	KRP3_COLLI	0.8	61	
Q9psv3	columba liv	100	KRP4_COLLI	0.8	62	
PL2707	xenopus lae	106	INS2_XENLA	0.8	c	
Q67295	aquifex aeo	106	Y094_HAETN	0.8	c	
P43939	haemophilus	106	Y094_HAETN	0.8	c	
P23460	rhodobacter	118	LHG2_RHOCA	0.8	c	
Q9ju11	neisseria m	119	CRCB_NEIMA	0.8	c	
Q91z96	neisseria m	119	CRCB_NEIMB	0.8	c	
Q9ck17	pasteurella	128	CRCB_PASMU	0.8	c	
P36090	thunnus obe	133	ATP6_THUOB	0.8	c	
P28164	chlamydia p	134	SRP_CHLPS	0.8	c	
P53716	candida alb	134	UBJ5_CANAL	0.8	c	
Q920h1	mus musculus	139	UGR1_MOUSE	0.8	c	
P06638	xenopus lae	141	HBA5_XENLA	0.8	c	
Q9ug74	homo sapien	144	PSG8_HUMAN	0.8	c	
Q30129	archaeoglob	153	PYR1_ARCFU	0.8	c	
Q8cf17	vibrio para	153	PYR1_VIBPA	0.8	c	
Q8dcf7	vibrio vuln	153	PYR1_VIBVU	0.8	c	
P46491	haemophilus	160	CRCB_HAETN	0.8	c	
Q9zds3	rickettsia	163	Y252_RICPR	0.8	c	
Q58649	methanococ	167	YCS1_METJA	0.8	c	
P18720	xenopus lae	168	ZG42_XENLA	0.8	c	
Q01548	hordium vul	170	PER2_HORVU	0.8	c	
PL7251	gallus gall	176	PTHR_CHICK	0.8	c	
P33467	feline ente	176	V6B_CVPE3	0.8	c	
Q8rae2	thermoanaer	185	EPF_THETN	0.8	c	
Q9vdx3	aeropyrum p	188	Y816_AERPE	0.8	c	
P56559	homo sapien	192	ARL7_HUMAN	0.8	c	
P55777	paramesium	193	NUZM_PARTE	0.8	c	
P39421	bacterioph	200	MODB_BPT4	0.8	c	
Q9ksw7	vibrio chol	210	HIS2_VIBCH	0.8	c	

c 92 7 0.8 210 1 HIS7 MYCLE
 c 93 7 0.8 211 1 ISPD CHLPN
 c 94 7 0.8 211 1 Y290_WETJA
 c 95 7 0.8 217 1 EVGI_HUMAN
 c 96 7 0.8 219 1 ATP6_ONCMA
 c 97 7 0.8 223 1 ATP6_ONCMY
 c 98 7 0.8 226 1 ATP6_XENLA
 c 99 7 0.8 227 1 ATP6_SALSA
 c 100 7 0.8 228 1 NEUA_HAEIN
 Q9x7b9 mycobacteri
 Q9z7x5 chlamydia p
 Q57738 methanococ
 Q9bze7 homo sapien
 Q36964 oncorhynchu
 P48178 oncorhynchu
 P00849 xenopus lae
 Q35920 salmo salar
 Q57140 haemophilus

ALIGNMENTS

RESULT 1

ID SM3A_HUMAN STANDARD; PRT; 771 AA.
 AC Q14563;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;

RX MEDLINE=94094332; PubMed=8269517;
 RA Kolodkin A.L., Matthes D.J., Goodman C.S.;

RT "The semaphorin genes encode a family of transmembrane and secreted
 RT growth cone guidance molecules.";
 RL Cell 75:1389-1399(1993).
 RN [2]

RP SEQUENCE OF 1-37 FROM N.A.
 RA Woessner J., Minx P., Hinds K., Strommatt C.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE OF 39-182 FROM N.A.
 RA Rohlfing T., Tin-Wollam A.M., Duckels G.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH

CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
 CC NEUROPILIN-1/PLEXIN-1 (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY

CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L26081; AAA65938.1; -;
 DR EMBL; AC004451; -; NOT_ANNOTATED_CDS.

DR EMBL; AC004848; AAC78622.1; -;
 DR PIR; D49423; D49423.

DR Genew; HGNC:10723; SEMA3A.
 DR MIM; 603961; -;

DR GO; GO:0005576; C:extracellular; TAS.
 DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.

DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 771
 FT DOMAIN 240 538
 FT DOMAIN 580 664
 FT DOMAIN 727 769
 FT DISULFD 649 722
 FT CARBOHYD 53 53
 FT CARBOHYD 125 125
 FT CARBOHYD 590 590
 SQ SEQUENCE 771 AA; 88889 MW; 9985F8D3EAED8456 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_HUMAN (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTGTCTCTTTCTGGGAGTATTACTACAGCAAGCA 259
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThAlaAla 20
 QY 260 AACTATCAGATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTCAAGAATAATGTG 319
 Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCACAATGTGATCAGTCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCCTTC 379
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGTGATGAGAACGGAGTAGGCTGTATCTTGGAGCAAGGATCATATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGTTAATATACAGGATTTTCAAAGATTGTGTGCCAGTATCTTACACCAAGA 499
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTGAAAGAAATGTCTAATTTTCATCAAGTA 559
 Db 101 AspGluCysLysTrpAlaGlyLysAspLeuLysGluCysAlaAsnPhelLysVal 120
 QY 560 CTTAAGGCATATAATACAGACTCACTTGTACGCCTGTGGAACGGGGCTTTTCATCCAATT 619
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTAATAATGGACATCATCTCGAGGACAATATTTTAACTGTGAGAACTCA 679
 Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAACACGGCTGGGAAGAGTCCATATGACCCCTTAAGCTGCTCAGACGATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThAlaSerLeu 180
 QY 740 TTAATAGATGGAGAAATATATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGACTTGTCT 799
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCGAACTCTTGGCCACCACCCCAATCAGGACAGACGACATGATTCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 850 CTCATGATCCAAAGTTTCATTAGTCCGCCACCTCATCTCAGAGAGTGACAATCTCGAAGAT 919
 Db 221 LeuAsnAspProLysPhePheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATCTTTTCTCCGTCGAAATGCAATAGATGGAGACACTCTCGGAAAAGCT 979
 Db 241 AspylsvaltyrPhePhePheargGluAsnAlaIleAspGlyGluHiserGlylysAla 260
 QY 980 ACTCAGCGTAGAATAGTGCAGATATGCAAGATGACTTTGGAGGCGCACAGAAGTCTGGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnileCyslysAsnAspPheGlyGlyHisargSerLeuVal 280
 QY 1040 AATAAATGGACAACTTCTCCTCAAGCTCGTCTGATTTGCTCAGTCCCGAGGCCAAATGGC 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTGATGACTGACAGATGATTCCTAATGCACTTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCAGTTGTATATGGAGTGTGTTACGACTTCCAGTAACTTTTCAAGGGATCAGCCGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGAGTGTAGAGAGGGTGTCTTGGTCCATATGCCACAGGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCAACTATCAATGGTGCTTATCAAGGAGAGTCCCTATCCACGCCCGAGACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCTCCAGCAAACTTTGGTGGTTTTCACCTCTCAAAAGGACCTTCTCGATGATGTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGCAAGAGTCATCCAGCCCATGATCAATCAGTGTTCCTATCAAACTCGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACCGATGTAATATCAATTTACAAATTTGCTAGACCGGTGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGACACATGATGATGTTTATCGGAACACAGATGTTGGACCGTCTTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 FTAGTTTCAATTCCTAAGGAGACTTGTATGATTTAGAGAGGTTTCTCTCGAAGAAATG 1639
 Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
 QY 1640 ACAGTTTTCGGAAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGTTTCAACGGCTGGGTTGCTCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
 Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGAAGAGCGTGTGCTGAGTGTGCTCCCGAGACCTTTACTGTGCTTGGATGGTTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGACGCCACAGACGACAGATATAGAAT 1879
 Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
 QY 1880 GGAGACCCACTGACTGCTTTCAGACTTTACACCATGATATACCATGGCCACAGCCCT 1939
 Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGAAATGCAGTCCGAG 1999

Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATAGAGAGCGCAAGAGAGAGAG 2059
 Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgGlnGluArgGlyGluGlu 620
 QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGAGTCTACAA 2119
 Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGATTTCAGCAATTTACTCTGCGATCGCGTGGAAACATGGGTTTCATCAAACTCTT 2179
 Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGGTAAACCTCGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATAAAG 2239
 Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
 QY 2240 GATCATGGAGATGCTCTAAGACCAAAAGAAATGCTCAATAGCATGACACCTAGCCAGAG 2299
 Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGGTACAGAGACTTCATGCGCTCATCAACACCCCAATCTCAACACCATGATGAG 2359
 Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
 QY 2360 TTCTGTCAACAAAGTTTCGAAAAAGGACCGAAAAACAACTCGGCAAGGCCAGGACATACC 2419
 Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
 QY 2420 CCAGGGAACAGTAACTAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
 Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
 QY 2480 ACCACAGAAATTTAGAGGGGACCCAGGAGTGTC 2512
 Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
 SM3A_MOUSE STANDARD; PRT; 772 AA.
 ID SM3A_MOUSE 008655; Q62180; Q62215;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D) (Sema D).
 DE SEMA3A OR SEMAD OR SEMD.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=NMRI; TISSUE=Embryo;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Puschel A.W., Adams R.H., Betz H.;
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-948(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97470885; PubMed=9331345;
 RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
 RA Yagi T.;
 RT "Disruption of semaphorin III/D gene causes severe abnormality in
 peripheral nerve projection.";
 RL Neuron 19:519-530(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kimura T., Fishman M.C.;
 RT "cDNA sequence of mouse collapsin/semaphorin III.";

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X95286; CAA64607.1; --
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 772 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 577 665 IG-LIKE C2-TYPE.
 FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 650 723 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 772 AA; 88808 MW; 240907812FF9P2D2 CRC64;

Alignment Scores:
 Pred. No.: 2.04e-89 Length: 772
 Score: 92.00 Matches: 92
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.28% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_RAT (1-772)

QY 1169 GTATGAGTGTTTACGACTTCAGTAACATTTTCAAGGATCAGCGTGATGAT 1228
 |||||
 Db 324 ValTyrGlyValPheThrSerSerAsnIlePheGlySerAlaValCysMetTyr 343
 |||||
 QY 1229 AGCATGAGTGTCAGAGGGTGTCTTCATGTCATATGCCAGGATGGACCCAAAC 1288
 |||||
 Db 344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
 |||||
 QY 1289 TATCATGGTGCTTATCAAGAGAGTCCCTATCCAGCGCAGGAACTTGTCCAGC 1348
 |||||
 Db 364 TyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
 |||||
 QY 1349 AAAACATTTGGTGTGGTTCCTACAAAGACCTTCTGTATGATGATGATTAACCTTGA 1408
 |||||
 Db 384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
 |||||
 QY 1409 AGAAGTCATCCAGCATGCTACAATCCAGTGTCTCT 1444
 |||||
 Db 404 ArgSerHisProAlaMetTyrAsnProValPhePro 415
 |||||

RESULT 4

SM3A_CHICK
 ID SM3A_CHICK STANDARD; PRT; 772 AA.
 AC Q90607;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Semaphorin 3A precursor (Collapsin-1) (COLL-1).
 GN SEM3A OR COLL1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

NCBI_TaxID=9031;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
 RP 666-680.
 RC TISSUE=Brain;
 RX MEDLINE=94006554; PubMed=8402908;
 RA Luo Y., Raible D., Raper J.A.;
 RT "Collapsin: a protein in brain that induces the collapse and paralysis
 of neuronal growth cones.";
 RL Cell 75:217-227(1993).
 CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROPILIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
 AND MUSCLE, MODERATE LEVELS IN LUNG, BURSA, AND HEART AND
 CC VIRTUALLY ABSENT IN LIVER. COLLAPSN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; U02528; AAC59638.1; --
 PIR; A49069; A49069.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 772 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 576 665 IG-LIKE C2-TYPE.
 FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 650 723 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 772 AA; 88867 MW; E91E09DE0CC940AC CRC64;

Alignment Scores:

Pred. No.: 5.33e-53 Length: 772
 Score: 58.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.48% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_CHICK (1-772)

QY 971 GGAAGCTACTCAGCTAGATAGTGCATATGCAAGATGACTTTTGAGGSCACAGA 1030
 |||||
 Db 258 GlyylsalatrHisAlaArgIleGlyIleCysLysAsnAspPheGlyGlyHisArg 277
 |||||
 QY 1031 AGTCTGTGTAATAATGGAACAACATCTCTCAAAGCTCGTCTGATTGCTCAGTGCCAGGT 1090

Db 278 SerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGly 297
 QY 1091 CCAATGCGATTCACACTATTTGATGAAGTTCAGGATGTTATCTATGAAC 1144
 Db 298 ProAsnGlyLeuAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 315

RESULT 5

SZ1A_BRARE STANDARD; PRT; 860 AA.
 ID SZ1A_BRARE
 AC Q9W701;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin Z1A precursor (Semaphorin 1A) (Sema-Z1A).
 GN SEMAZ1A OR SEMA3AA.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99313409; PubMed=10386838;
 RA Yee C.S., Chandrasekhar A., Halloran M.C., Shoji W., Warren J.T.,
 RA Kuwada J.Y.;
 RT "Molecular cloning, expression, and activity of zebrafish semaphorin
 Z1a.";
 RL Brain Res. Bull. 48:581-593(1999).
 CC -!- FUNCTION: MAY INFLUENCE OUTGROWTH BY A VARIETY OF GROWTH CONES
 CC INCLUDING THOSE OF THE POSTERIOR LATERAL LINE GANGLION.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN HIGHLY SPECIFIC PATTERNS WITHIN
 CC THE DEVELOPING EMBRYO.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF086761; AAD43964.1; -;
 CC ZFIN; ZDB-GENE-991209-3; sema3aa.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG-LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 17
 CC CHAIN 18 860
 CC DOMAIN 241 539
 CC DOMAIN 579 668
 CC DOMAIN 722 858
 CC DISULFID 652 717
 CC CARBOHYD 53 53
 CC CARBOHYD 126 126
 CC CARBOHYD 593 593
 CC SEQUENCE 860 AA; 97263 MW; 5FD4C12194F5165C CRC64;

Alignment Scores:
 Pred. No.: 2,28e-28 Length: 860
 Score: 35.00 Matches: 35
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SZ1A_BRARE (1-860)

QY 1307 CAAGAAGAGTCCCTATCCACGGCCAGAACTGTCCACGAAACATTGGTGTGTTT 1366
 Db 371 GlnGlyArgValProTyrProArgProGlyThrCysProSerLysThrPheGlyGlyPhe 390
 QY 1367 GACTCTACAAAGGACCTTCCTGATGATGTTATACCTTTGCAAGA 1411
 Db 391 AspSerThrLysAspLeuProAspValIleThrPheAlaArg 405

RESULT 6

SZ1B_BRARE STANDARD; PRT; 778 AA.
 ID SZ1B_BRARE
 AC Q9W686;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin Z1B precursor (Semaphorin 1B) (Sema-Z1B).
 GN SEMAZ1B OR SEMA3AB.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425174; PubMed=10495275;
 RA Roos M., Schachner M., Bernhardt R.R.;
 RT "Zebrafish semaphorin Z1b inhibits growing motor axons in vivo.";
 RL Mech. Dev. 87:103-117(1999).
 CC -!- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
 CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
 CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
 CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
 CC VENTRALLY EXTENDING MOTOR AXONS.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF083382; AAD28103.1; -;
 CC ZFIN; ZDB-GENE-991209-6; sema3ab.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin-repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG-LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.

CC SIGNAL 1 17
 CC CHAIN 18 860
 CC DOMAIN 241 539
 CC DOMAIN 579 668
 CC DOMAIN 722 858
 CC DISULFID 652 717
 CC CARBOHYD 53 53
 CC CARBOHYD 126 126
 CC CARBOHYD 593 593
 CC SEQUENCE 860 AA; 97263 MW; 5FD4C12194F5165C CRC64;

KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.
FT SIGNAL; 1 17 POTENTIAL.
FT CHAIN 18 778 SEMAPHORIN 21B.
FT DOMAIN 241 539 SEMA.
FT DOMAIN 579 668 IG-LIKE C2-TYPE.
FT DOMAIN 721 776 ARG/LYS-RICH (BASIC).
FT DISULFID 652 716 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 778 AA; 88904 MW; 4D36F432A2E1895 CRC64;

Alignment Scores:
Pred. No.: 1.39e-16 Length: 778
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.68% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x S21B_BRARE (1-778)
QY 1823 TGTTCGCTATTTCCCACTGCAAGAGACGACACAGACAGATATAAGAAATGGA 1882
DB 543 CysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsnGly 562

QY 1883 GACCACTGACT 1894
DB 563 AspProLeuThr 566

RESULT 7
SM3B_MOUSE
ID SM3B_MOUSE STANDARD; PRT; 748 AA.
AC Q62177;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin A) (Sema A).
GN SEMA3B OR SEMA OR SEMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; PubMed=7748561;
RA Puechel A.W., Adams R.H., Bez H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X85990; CAA59982.1; -.
CC PIR; I48744; I48744.
CC MGD; MGI:107561; Sema3b.
CC InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR01627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
Developmental protein; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 748 SEMAPHORIN 3B.
FT DOMAIN 55 496 SEMA.
FT DOMAIN 561 659 IG-LIKE C2-TYPE.
FT DOMAIN 698 702 POLY-GLY.
FT DOMAIN 723 743 ARG-RICH (BASIC).
FT DISULFID 643 709 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 748 AA; 82894 MW; 1866B7D2397C9305 CRC64;

Alignment Scores:
Pred. No.: 2.68e-12 Length: 748
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3B_MOUSE (1-748)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGATGTTCTGCATGT 1825
DB 521 AlaCysAlaGluCysCysLeuAlaArgAspProIrrCysAlaIrrPaspGlySerAlaCys 540

RESULT 8
SM22_BRARE
ID SM22_BRARE STANDARD; PRT; 764 AA.
AC Q9W6G6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin Z2 precursor (Semaphorin 2) (Sema-22).
GN SEMA22 OR SEMA2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99112778; PubMed=9915572;
RA Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Raper J.A.,
Kuwaada J.Y.;
RT "Analysis of a Zebrafish semaphorin reveals potential functions in
vivo.";
RL Dev. Dyn. 214:13-25(1999).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE GUIDANCE OF SEVERAL AXON
PATHWAYS.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN A DYNAMIC AND RESTRICTED PATTERN
DURING THE PERIOD OF AXON OUTGROWTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X85990; CAA59982.1; -.
CC PIR; I48744; I48744.
CC MGD; MGI:107561; Sema3b.
CC InterPro; IPR007110; Ig-like.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL; AF124485; AAD21310.1; -.
CC ZFIN; ZDB-GENE-990715-2; sema2.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00409; Ig; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 41
FT CHAIN 42 764
FT DOMAIN 261 559
FT DOMAIN 661 740
FT DOMAIN 741 762
FT DISULFID 668 733
FT CARBOHYD 143 143
FT CARBOHYD 490 490
FT CARBOHYD 610 610
SQ SEQUENCE 764 AA; 87859 MW; A3ED95C2C479D7AE CRC64;

Alignment Scores:
Pred. No.: 5,16e-08 Length: 764
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.79% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SMZ2_BARE (1-764)

QY 1769 TGTGCTGAGTGTGCTCCGACACCTTACTGCTGGATGGT 1816
Db 545 CysAlaGluCysCysLeuAlaArgAspProtyrCysAlaTrpAspGly 560

RESULT 9
ID SM3B_HUMAN STANDARD; PRT; 749 AA.
AC Q13214; Q8TDV7; Q93018;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin V) (Sema V) (Sema A(V)).
GN SEMA3B OR SEMA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Dante M., Wamsley P.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Koyama N.;
RT "Semaphorin 3B (SEMA3B) cDNA.";
RN [4]
RP Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RX VARIANTS NSCLC CVS-348; H18-397 AND ILE-415.
RX MEDLINE=20535986; PubMed=11085536;
RA Lerman M.I., Minna J.D.;

```

```

RT "The 630-kb lung cancer homozygous deletion region on human chromosome
RT 3p21.3: identification and evaluation of the resident candidate tumor
RT suppressor genes."
RL Cancer Res. 60:6116-6133(2000).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
CC ENDOPLASMIC RETICULUM.
CC -!- TISSUE SPECIFICITY: Expressed abundantly but differentially in a
CC variety of neural and nonneural tissues.
CC -!- DISEASE: Defects in SEMA3B are found in non-small cell lung cancer
CC (NSCLC) cell lines.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Alignment Scores:
 Pred. No.: 6.08e-07 Length: 749
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0

DB: 1 0 Gaps: 0

US-09-774-490-1 (1-2709) x SM3B_HUMAN (1-749)

QY 1028 AGAAGTCGGTGAATAAATGGAACATTCCTCAAGCTCGTCTG 1072
|||||

Db 277 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 291
|||||

RESULT 10

SM3C_CHICK

ID SM3C_CHICK STANDARD; PRT; 751 AA.

AC 042336; O90664;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 3C precursor (Collapsin-3) (COLL-3).

GN SEMA3C OR COLL3.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;

RX MEDLINE=95329269; PubMed=7605628;

RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;

RT "A family of molecules related to collapsin in the embryonic chick nervous system";

RL Neuron 14:1131-1140(1995).

CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: COLLAPSIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT DISTINCT AXON TRACTS.

CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY THIRD OF THE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF022946; AAB80951.1; --

CC EMBL; U28241; AAB86897.1; --

CC InterPro; IPR007110; IG-like.

CC InterPro; IPR003599; IG.

CC InterPro; IPR003659; Plexin-like.

CC Pfam; PF01403; Sema; 1.

CC SMART; SM00409; IG; 1.

CC SMART; SM00423; PSI; 1.

CC SMART; SM00630; Sema; 1.

CC PROSITE; PS50835; IG LIKE; 1.

KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 751 SEMAPHORIN 3C.

FT DOMAIN 237 535 SEMA.

FT DOMAIN 571 655 IG-LIKE C2-TYPE.

FT DOMAIN 718 749 ARG/LYS-RICH (BASIC).

FT DISULFID 643 709 BY SIMILARITY.

FT CARBOHYD 81 81 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 123 123 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 585 585 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 238 238 A -> D (IN REF. 2).

FT CONFLICT 359 359 P -> S (IN REF. 2).

FT CONFLICT 404 404 H -> D (IN REF. 2).

SQ SEQUENCE 751 AA; 85433 MW; DB686687D21BD5D8 CRC64;

Alignment Scores:

Pred. No.: 6.08e-07 Length: 751

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.68% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3C_CHICK (1-751)

QY 1028 AGAAGTCGGTGAATAAATGGAACATTCCTCAAGCTCGTCTG 1072
|||||

Db 274 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 288
|||||

RESULT 11

SM3C_HUMAN

ID SM3C_HUMAN STANDARD; PRT; 751 AA.

AC Q99985;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 3C precursor (Semaphorin E) (Sema E).

GN SEMA3C OR SEMAE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=98070817; PubMed=9405678;

RA Yamada T., Endo R., Gotoh M., Hirohashi S.;

RT "Identification of semaphorin E as a non-MDR drug resistance gene of human cancers";

RL Proc. Natl. Acad. Sci. U.S.A. 94:14713-14718(1997).

CC -!- FUNCTION: MAY BE INVOLVED IN DIVERSE CELL SURVIVAL MECHANISMS.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: EXPRESSED INTENSELY IN THE HEART, SKELETAL MUSCLE, COLON, SMALL INTESTINE, OVARY, TESTIS, AND PROSTATE. PAINT EXPRESSION UBQUITOUSLY AMONG OTHER ORGANS, INCLUDING BRAIN.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB000220; BAA32398.1; --

CC Genbank; HGNC:10725; SEMA3C.

CC MIM; 602645; --

CC GO; GO:0008151; P:cell growth and/or maintenance; TAS.

CC GO; GO:0009315; P:drug resistance; TAS.

CC GO; GO:0006955; P:immune response; TAS.

CC InterPro; IPR007110; IG-like.

CC InterPro; IPR003599; IG.


```

CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC004957; RAC3081.1; ALT_SEQ.
CC Genew; HGNC:10726; SEMA3D.
CC InterPro; IPR007110; Ig-Like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001827; Sema.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 39
FT CHAIN 40 777 SEMAPHORIN 3D.
FT DOMAIN 257 555 SEMA.
FT DOMAIN 592 680 IG-LIKE C2-TYPE.
FT DOMAIN 739 773 ARG/LYS-RICH (BASIC).
FT DISULFID 665 731 BY SIMILARITY.
FT CARBOHYD 139 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 724 724 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 777 AA; 89651 MW; 3F7B0D7AF50F53BD CRC64;

Alignment Scores:
Pred. No.: 7,14e-06 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3D_HUMAN (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuLysCysSer 311

RESULT 14
SM3D_CHICK
ID SM3D_CHICK STANDARD; PRT; 761 AA.
AC Q90663;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3D precursor (Collapsin-2) (COLL-2).
GN SEMA3D OR COLL2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95329269; PubMed=7605628;
RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT "A family of molecules related to collapsin in the embryonic chick
RT nervous system.";
RL Neuron 14:1131-1140(1995).
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH

```

```

CC CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
CC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: DEVELOPING SPINAL CORD AND DEVELOPING VISUAL
CC SYSTEM. COLLAPSIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT
CC DISTINCT AXON TRACTS.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U28240; AAA86896.1; -.
CC InterPro; IPR007110; Ig-Like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 761 SEMAPHORIN 3D.
FT DOMAIN 245 543 SEMA.
FT DOMAIN 552 670 IG-LIKE C2-TYPE.
FT DOMAIN 727 757 ARG/LYS-RICH (BASIC).
FT DISULFID 653 719 BY SIMILARITY.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 761 AA; 87300 MW; 3E09AE3DBA53F46B CRC64;

Alignment Scores:
Pred. No.: 8.43e-05 Length: 761
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3D_CHICK (1-761)

QY 1778 TGTTGCTGCGCCGAGACCTTACTGTGTTGGATGGT 1816
Db 532 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544

RESULT 15
SM3E_HUMAN
ID SM3E_HUMAN STANDARD; PRT; 775 AA.
AC O15041;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3E precursor.
GN SEMA3E OR KIAA0331.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Brain;

```



```
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [2]
RP SEQUENCE OF 1-38 FROM N.A.
RA Holmes A., Johnson D., Elliot G., Burkhardt J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB002329; BAA20789.1; -.
DR EMBL; AC004954; AAC69513.1; -.
DR GenBank; HGNC:10727; SEMA3E
DR GO; GO:0005613; C:extracellular space; TAS.
DR GO; GO:0007048; P:oncogenesis; TAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 775 SEMAPHORIN 3E.
FT DOMAIN 241 540 SEMA.
FT DOMAIN 581 669
FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).
FT DISULFD 654 729
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 775 AA; 89227 MW; CD6079C1AB48F779 CRC64;

Alignment Scores:
Pred. No.: 8,41e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_HUMAN (1-775)
QY 1778 TGTTCCTCGCCGAGACCTTACTGTGCTGGATGGT 1816
Db 529 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 16
SM3E_MOUSE STANDARD; PRT; 775 AA.
ID SM3E_MOUSE
AC P70275; O09078; O09079;
```

```
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3E precursor (Semaphorin H) (Sema H).
GN SEMA3E OR SEMAR OR SEMH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RX MEDLINE=98175564; PubMed=9515811;
RA Christensen C.R.L., Klingelhofer J., Tarabykina S., Hulgaard E.F.,
RA Kramarov D., Lukanidin E.;
RT "Transcription of a novel mouse semaphorin gene, M-semaH, correlates
RT with the metastatic ability of mouse tumor cell lines.";
RL Cancer Res. 58:1238-1244(1998).
RN [2]
RP REVISIONS
RA Christensen C.R.L.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: DEVELOPING LUNGS, DEVELOPING SKELETAL
CC ELEMENTS, AND VENTRAL HORNS OF THE DEVELOPING NEURAL TUBE.
CC CORRELATES POSITIVELY WITH TUMOR PROGRESSION.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z80941; CAB02590.1; -.
DR EMBL; Z93947; CAB07987.1; ALT SEQ.
DR EMBL; Z93948; CAB07988.1; ALT_SEQ.
DR MGD; MGI:1340034; Sema3e.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 775 SEMAPHORIN 3E.
FT DOMAIN 241 540 SEMA.
FT DOMAIN 581 669
FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).
FT DISULFD 654 729
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 775 AA; 89503 MW; BA41690093D28F40 CRC64;

Alignment Scores:
Pred. No.: 8,41e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
```

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_MOUSE (1-775)
 QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGGATGCT 1816
 DB 529 CysCysLeuAlaArgAspProTyrcysAlaTrpAspGly 541

RESULT 17
 SM3E_CHICK
 ID SM3E_CHICK STANDARD; PRT; 785 AA.
 AC 042237; Q90666;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3E precursor (Collapsin-5) (COLL-5).
 GN SEMA3E OR COLL5
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=97470887; PubMed=9331347;
 RA Feiner L., Koppel A.M., Kobayashi H., Raper J.A.;
 RT "Secreted chick semaphorins bind recombinant neuropilin with similar
 RT affinities but bind different subsets of neurons in situ."
 RL Neuron 19:539-545(1997).
 RN [2]
 RP SEQUENCE OF 244-543 FROM N.A.
 RX MEDLINE=95329269; PubMed=7605628;
 RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
 RT "A family of molecules related to collapsin in the embryonic chick
 RT nervous system."
 RL Neuron 14:1131-1140(1995).
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
 CC NEURONAL POPULATIONS. BINDS TO NEUROFILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: COLLAPSIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF022947; AAB80952.1; -;
 DR EMBL; U28243; AAB86899.1; -;
 DR InterPro; IPR003659; Plexin-like.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 1 25 SEMAPHORIN 3E.
 FT DOMAIN 26 785 SEMA.
 FT DOMAIN 244 544 SEMA.
 FT DOMAIN 651 740 IG-LIKE C2-TYPE.
 FT DOMAIN 741 780 ARG/LYS-RICH (BASIC).

FT DISULFID 658 733 BY SIMILARITY.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 600 600 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 246 246 N -> D (IN REF. 2).
 FT CONFLICT 248 248 V -> I (IN REF. 2).
 FT CONFLICT 250 250 L -> F (IN REF. 2).
 SQ SEQUENCE 785 AA; 90978 MW; E551EBF717630632 CRC64;

Alignment Scores:
 Pred. No.: 8.41e-05 Length: 785
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_CHICK (1-785)
 QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGGATGCT 1816
 DB 533 CysCysLeuAlaArgAspProTyrcysAlaTrpAspGly 545

RESULT 18
 SM3F_HUMAN
 ID SM3F_HUMAN STANDARD; PRT; 785 AA.
 AC Q13275; Q13274; Q13372; Q15704;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV) (Sema III/F).
 GN SEMA3F.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96226360; PubMed=8649831;
 RA Roche J., Boldog F., Robinson M., Robinson L., Varella-Garcia M.,
 RA Swanton M., Waggoner B., Fishel R., Franklin W., Gemmill R.,
 RA Drabkin H.;
 RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
 RT human semaphorin."
 RL Oncogene 12:1289-1297(1996).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT MET-503.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96230324; PubMed=8786119;
 RA Xiang R.-H., Hensel C.H., Garcia D.K., Carlson H.C., Kok K.,
 RA Daly M.C., Korbacher K., van den Berg A., Veldhuis P., Buys C.H.C.M.,
 RA Naylor S.L.;
 RT "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
 RT 3p21, a region deleted in lung cancer."
 RL Genomics 32:39-48(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Nelson J., Biewald T.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 394-436 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96210603; PubMed=8633026;
 RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
 RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
 RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
 RT cancer deletion region and demonstrate distinct expression patterns."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
 CC -!- FUNCTION: MAY PLAY A ROLE IN CELL MOTILITY AND CELL ADHESION.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
 CC VARIETY OF NEURAL AND NONNEURAL TISSUES. THERE IS HIGH EXPRESSION

CC IN MAMMARY GLAND, KIDNEY, FETAL BRAIN, AND LUNG AND LOWER
 CC EXPRESSION IN HEART AND LIVER.
 CC -!- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 10.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL: U33920; AAC50568.1; --
 CC EMBL: U32876; AAB18276.1; --
 CC EMBL: AC000063; AAB46344.1; --
 CC EMBL: U32171; AAB06011.1; --
 CC EMBL: U32172; AAB06012.1; --
 CC Genew; HGNC:10728; SEMA3F.

CC MIM: 601124; --
 CC GO: GO:0005615; C:extracellular space; TAS.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG-LIKE; 1.

CC Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Polymorphism.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 785 SEMAPHORIN 3F.
 FT DOMAIN 272 569 SEMA.
 FT DOMAIN 605 690 IG-LIKE C2-TYPE.
 FT DOMAIN 758 779 ARG/LYS-RICH (BASIC).
 FT DISULFID 678 746 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 474 474 A -> G (in dbSNP:1046955).
 FT L -> M.
 FT VARIANT 503 503 /FTID=VAR_011820.
 FT L -> W.
 FT VARIANT 503 503 /FTID=VAR_008855.
 FT CONFLICT 153 183 MISSING (IN REF. 2).
 FT CONFLICT 270 270 MISSING (IN REF. 2).
 FT CONFLICT 473 473 A -> S (IN REF. 2).
 SQ SEQUENCE 785 AA; 88381 MW; FE3FC796BEC1608E CRC64;

Alignment Scores:
 Pred. No.: 8,41e-05 Length: 785
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3F_HUMAN (1-785)
 QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGGATGCT 1816
 DB 558 CysCysLeuAlaArgAspProTrpCysAlaTrpAspGly 570

RESULT 19
 SM3F_MOUSE
 ID SM3F_MOUSE STANDARD; PRT; 785 AA.
 AC O88632; O88633;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV).
 GN SEMA3F.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99090943; PubMed=9875739;
 RA Eckhardt F., Meyerhans A.;
 RT "Cloning and expression pattern of a murine semaphorin homologous to
 RL H-sema IV.";
 RL NeuroReport 9:3975-3979(1998).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=B;
 CC IsoId=O88632-1; Sequence=Displayed;
 CC Name=A;
 CC IsoId=O88632-2; Sequence=VSP_006042;
 CC -!- TISSUE SPECIFICITY: EXPRESSED UBIQUITOUSLY IN ADULTHOOD. DURING
 CC EMBRYOGENESIS, EXPRESSED IN SUBREGIONS OF THE CENTRAL NERVOUS
 CC SYSTEM AND VARIOUS OTHER TISSUES LIKE SKIN, KIDNEY, LUNG AND
 CC INTESTINE.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF080090; AAC28108.1; --
 CC EMBL: AF080091; AAC28109.1; --
 CC MGD; MGI:1096347; Sema3f.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003598; IG_C2.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00408; IGC2; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG-LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 785 SEMAPHORIN 3F.
 FT DOMAIN 272 569 SEMA.
 FT DOMAIN 605 695 IG-LIKE C2-TYPE.
 FT DOMAIN 758 779 ARG/LYS-RICH (BASIC).
 FT DOMAIN 715 718 POLY-PRO.
 FT DISULFID 678 746 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 153 183 Missing (in isoform A).
 FT /FTID=VSP_006042;
 SQ SEQUENCE 785 AA; 88493 MW; 89DBA7A08D02A72E CRC64;

Alignment Scores:
 Pred. No.: 8,41e-05 Length: 785
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

```

US-09-774-490-1 (1-2709) x SM3F_MOUSE (1-785)
QY 1778 TGTTCCTCGCCGAGACCTTACTGTGCTTGGGATGGT 1816
      |||||||
Db 558 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 570

RESULT 20
SM4C_MOUSE
ID SM4C_MOUSE STANDARD; PRT; 834 AA.
AC O64151;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4C precursor (Semaphorin 1) (Sema 1) (Semaphorin C-like 1)
DE (M-Sema F).
GN SEMA4C OR SEMA1 OR SEMACL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Neonatal brain;
RX MEDLINE=95385809; PubMed=7656991;
RA Inagaki S., Furuyama T., Iwahashi Y.;
RT "Identification of a member of mouse semaphorin family.";
RL FEBS Lett. 370:269-272(1995).
RN [2]
RP INTERACTION WITH GIPC.
RX MEDLINE=99253973; PubMed=10318831;
RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT "A PDZ protein regulates the distribution of the transmembrane
semaphorin, M-SemF.";
RL J. Biol. Chem. 274:14137-14146(1999).
CC -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED WIDELY IN THE NERVOUS TISSUES DURING
DEVELOPMENT. PREDOMINANTLY IN THE DEVELOPING BRAIN AND SPINAL
CORD.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; S79463; AAB35184.1; -
CC PIR; S66498; S66498.
CC MGD; MGI:109252; Sema4c.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01437; PSI; 1.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS50835; IG LIKE; 1.
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 834 SEMAPHORIN 4C.
FT DOMAIN 21 664 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 665 685 POTENTIAL.

US-09-774-490-1 (1-2709) x SM4C_MOUSE (1-834)
QY 1043 AAATGGACACATTCCTCAAGCTCGTCTG 1072
      |||||||
Db 274 LysTrpThrPheLeuLysAlaArgLeu 283

RESULT 21
SM4D_CHICK
ID SM4D_CHICK STANDARD; PRT; 295 AA.
AC Q90655;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4D (Collapsin-4) (COLL-4) (Fragment).
GN SEMA4D OR COLL4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95329269; PubMed=7605628;
RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT "A family of molecules related to collapsin in the embryonic chick
nervous system.";
RL Neuron 14:1131-1140(1995).
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
NEURONAL POPULATIONS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U28242; AAA86898.1; -
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01437; PSI; 1.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00630; Sema; 1.
KW Multigene family; Neurogenesis; Developmental protein; Glycoprotein.

```

```
FT NON_TER 1 1 SEMA.
FT DOMAIN <1 255
FT PSI 275 >295
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 295 295
SQ SEQUENCE 295 AA; 33524 MW; 298AA49414C352F0A CRC64;

Alignment Scores:
Pred. No.: 1.76 Length: 295
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4D_CHICK (1-295)
OY 1784 CTCGCCGAGACCTTACTGTCTTGG 1810
Db 287 LeuAlaArgAspProTyfCysAlaTrp 295

RESULT 22
CIS_HUMAN
ID CIS_HUMAN STANDARD; PRT; 688 AA.
AC P09871; Q9UCU7; Q9UCU8; Q9UCU9; Q9UCV0; Q9UCV1; Q9UCV2; Q9UCV3;
AC Q9UCV4; Q9UCV5; Q9UM14;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C1s component precursor (EC 3.4.21.42) (C1 esterase).
GN C1s.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89017187; PubMed=2459702;
RA Kusunoto H., Hirose S., Salter J.-P., Hagen F.S., Kurachi K.;
RT "Human genes for complement components C1r and C1s in a close
tail-to-tail arrangement.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7307-7311(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88082788; PubMed=3500856;
RA McKinnon C.M., Carter P.E., Smyth S.J., Dunbar B., Fothergill J.E.;
RT "Molecular cloning of cDNA for human complement component C1s. The
complete amino acid sequence.";
RL Eur. J. Biochem. 169:547-553(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86163522; PubMed=2831944;
RA Tosi M., Duponchel C., Meo T., Julier C.;
RT "Complete cDNA sequence of human complement C1s and close physical
linkage of the homologous genes C1s and C1r.";
RL Biochemistry 26:8516-8524(1987).
RN [4]
RP SEQUENCE OF 1-329 FROM N.A.
RC TISSUE=Peripheral blood leukocytes;
RX MEDLINE=99008558; PubMed=9794427;
RA Endo Y., Takahashi M., Nakao M., Saiga H., Sekine H., Matsushita M.,
RA Nonaka M., Fujita T.;
RT "Two lineages of mannose-binding lectin-associated serine protease
(MASP) in vertebrates.";
RL J. Immunol. 161:4924-4930(1998).
RN [5]
RP SEQUENCE OF 291-688 FROM N.A.
RX MEDLINE=90040704; PubMed=2553984;
RA Tosi M., Duponchel C., Meo T., Couture-Tosi E.;
RT "Complement genes C1r and C1s feature an intronless serine protease
domain closely related to haptoglobin.";
```

```
RL J. Mol. Biol. 208:709-714(1989).
RN [6]
RP SEQUENCE OF 16-61; 168-219; 287-334 AND 384-445.
RX MEDLINE=86164350; PubMed=3007145;
RA Spycher S.E., Nick H., Rickli E.E.;
RT "Human complement component C1s. Partial sequence determination of
the heavy chain and identification of the peptide bond cleaved during
activation.";
RL Eur. J. Biochem. 156:49-57(1986).
RN [7]
RP SEQUENCE OF 438-500; 503-534; 542-601; 617-623 AND 626-656.
RX MEDLINE=84104122; PubMed=632661;
RA Carter P.E., Dunbar B., Fothergill J.E.;
RT "The serine proteinase chain of human complement component C1s.
Cyanogen bromide cleavage and N-terminal sequences of the
fragments.";
RL Biochem. J. 215:565-571(1983).
RN [8]
RP PARTIAL SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=91308095; PubMed=1854725;
RA Illy C., Thielens N.M., Gagnon J., Arlaud G.J.;
RT "Effect of lactoperoxidase-catalyzed iodination on the Ca(2+)-
dependent interactions of human C1s. Location of the iodination
sites.";
RL Biochemistry 30:7135-7141(1991).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=91175725; PubMed=2007122;
RA Hess D., Schaller J., Rickli E.E.;
RT "Identification of the disulfide bonds of human complement C1s.";
RL Biochemistry 30:2827-2833(1991).
RN [10]
RP PARTIAL SEQUENCE, AND 3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
RX MEDLINE=95298736; PubMed=779774;
RA Rossi V., Gaboriaud C., Lacroix M., Ulrich J., Fontecilla-Camps J.C.,
RA Gagnon J., Arlaud G.J.;
RT "Structure of the catalytic region of human complement protease C1s:
study by chemical cross-linking and three-dimensional homology
modeling.";
RL Biochemistry 34:7311-7321(1995).
RN [11]
RP DISEASE.
RX MEDLINE=21286517; PubMed=11390518;
RA Dragon-Durey M.-A., Quartier P., Fremaux-Bacchi V., Blouin J.,
RA de Barace C., Prieur A.-M., Weiss L., Fridman W.-H.;
RT "Molecular basis of a selective C1s deficiency associated with early
onset multiple autoimmune diseases.";
RL J. Immunol. 166:7612-7616(2001).
RN [12]
RP FUNCTION: C1s B CHAIN IS A SERINE PROTEASE THAT COMBINES WITH C1Q
AND C1s TO FORM C1, THE FIRST COMPONENT OF THE CLASSICAL PATHWAY
OF THE COMPLEMENT SYSTEM. C1r ACTIVATES C1s SO THAT IT CAN, IN
TURN, ACTIVATE C2 AND C4.
RN [13]
RP CATALYTIC ACTIVITY: Cleaves component C4 to C4a and C4b (Arg-|-Ala
bond), and component C2 to C2a and C2b (Lys(or Arg)-|-Lys bond).
RN [14]
RP SUBUNIT: C1s IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
AND S IN THE MOLAR RATIO OF 1:2:2. ACTIVATED C1s IS AN DISULFIDE-
LINKED HETERODIMER OF AN HEAVY CHAIN AND A LIGHT CHAIN.
RN [15]
RP DISEASE: Defects in C1s are the cause of selective C1s deficiency
[MIM:120580]; that is associated with early onset multiple
autoimmune diseases.
RN [16]
RP SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
RN [17]
RP SIMILARITY: Contains 2 Sushi (SCR) domains.
RN [18]
RP SIMILARITY: Contains 2 CUB domains.
RN [19]
RP SIMILARITY: Contains 1 EGF-like domain.
RN [20]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```

CC EMBL; X06596; CA29817.1; -.
DR EMBL; J04080; AAA51852.1; -.
DR EMBL; M18767; AAA51853.1; -.
DR EMBL; AB009076; BAA86864.1; -.
DR PDB; 1ELV; 14-MAR-01.
DR MEROPS; S01.193; -.
DR SWISS-2DPAGE; P09871; HUMAN.
DR Genew; HGNC:1247; CIS.
DR MIM; 120580; -.
DR GO; GO:0003816; F:complement component C1s activity; TAS.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001254; Ser protease Try.
DR InterPro; IPR000436; Sushi_SCR_CCF.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00084; sushi; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00022; EGF 1; FALSE NEG.
DR PROSITE; PS01186; EGF 2; FALSE NEG.
DR PROSITE; PS01187; EGF CA; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Complement pathway; Plasma; Glycoprotein; Serine protease; Hydrolase;
KW Hydroxylation; Sushi; Repeat; Signal; EGF-like domain;
KW Calcium-binding; Polymorphism; 3D-structure.
FT SIGNAL 1 15
FT CHAIN 16 437
FT CHAIN 438 688
FT DOMAIN 16 130
FT DOMAIN 131 172
FT DOMAIN 175 290
FT DOMAIN 293 355
FT DOMAIN 358 422
FT DOMAIN 438 688
FT ACT_SITE 475 475
FT ACT_SITE 529 529
FT ACT_SITE 632 632
FT MOD_RES 149 149
FT CARBOHYD 174 174
FT CARBOHYD 406 406
FT DISULFID 65 83
FT DISULFID 135 147
FT DISULFID 143 156
FT DISULFID 158 171
FT DISULFID 175 202
FT DISULFID 234 251
FT DISULFID 294 341
FT DISULFID 321 354
FT DISULFID 359 403
FT DISULFID 386 421
FT DISULFID 425 549
FT DISULFID 595 618
FT DISULFID 628 659
FT VARIANT 383 383
FT CONFLICT 294 294
FT CONFLICT 513 513
FT CONFLICT 573 573
FT CONFLICT 645 646
SQ SEQUENCE 688 AA; 76684 MW; 85522647A4C47205 CRC64;

INTERCHAIN.
R -> H (IN dbSNP:20573) .
/FTID=VAR_014565.
C -> K (IN REF. 6).
G -> GG (IN REF. 5).
T -> A (IN REF. 7).
TK -> GR (IN REF. 7).
SEQUENCE 688 AA; 76684 MW; 85522647A4C47205 CRC64;

Alignment Scores:
Pred. No.: 1.64 Length: 688
Score: 9.00 Matches: 9

```

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x CIS_HUMAN (1-688)

QY 151 TTGTGCGCCAGAGAGTTCACAAAT 125
Db 82 LeuCysGlyGlnArgSerSerAsnAsn 90

RESULT 23
SN4B MOUSE
ID SM4B MOUSE STANDARD; PRT; 782 AA.
AC Q62179;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
GN SEMA4B OR SEMAC OR SEMC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Brain;
RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RL Neuron 14:941-948 (1995).
RP [2]
RP INTERACTION WITH GIPC.
RX MEDLINE=95253973; PubMed=10318831;
RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT "A PDZ protein regulates the distribution of the transmembrane
RL semaphorin, M-SemF."
RX J. Biol. Chem. 274:14137-14146 (1999).
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X85992; CAA59984.1; -.
PIR; I48746; I48746.
MGD; MGI:107559; Sema4b.
InterPro; IPR003659; Plexin-like.
InterPro; IPR002165; Plexin_repeat.
InterPro; IPR001627; Sema.
Pfam; PF01437; PSI; 1.
Pfam; PF01403; Sema; 1.
SMART; SM00423; PSI; 1.
SMART; SM00630; Sema; 1.
Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
Developmental protein; Glycoprotein.
NON_TER 1
DOMAIN <1 662 EXTRACELLULAR (POTENTIAL).
TRANSMEM 663 683 POTENTIAL.
FT

```


Pred. No.: 1-61 Length: 832
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4B_HUMAN (1-832)
 QY 1784 CTCGCCGAGACCTTACTGCTTGG 1810
 |||||
 Db 532 LeuAlaArgProTyrCysAlaTrp 540
 |||||

RESULT 25
 SM4D MOUSE
 ID SM4D MOUSE STANDARD; PRT; 861 AA.
 AC Q09126;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4D precursor (Semaphorin J) (Sema J) (Semaphorin C-like 2) (M-Sema G).
 DE (M-Sema G).
 GN SEMA4D OR SEMAJ OR SEMA2L2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=97125976; PubMed=8969198;
 RA Fuyuyama T., Inagaki S., Kosugi A., Noda S., Saitoh S.-I., Ogata M.,
 RA Iwahashi Y., Miyazaki N., Hamada T., Tohyama M.;
 RT "Identification of a novel transmembrane semaphorin expressed on
 RT lymphocytes.";
 RL J Biol. Chem. 271:33376-33381(1996).
 CC -!- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL
 CC AS IN THE NERVOUS SYSTEM.
 CC -!- TISSUE SPECIFICITY: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN LYMPHOID TISSUES,
 CC ESPECIALLY IN THE THYMUS, AS WELL AS IN THE NERVOUS TISSUES.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U69535; AAC52964.1; -.
 CC MGD; MGI:109244; Sema4d.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00408; IGC2; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG_LIKE; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 CC Neurogenesis; Developmental protein; Glycoprotein.
 KW SIGNAL 1 23 POTENTIAL.
 FT SIGNAL 24 861 SEMAPHORIN 4D.
 FT CHAIN 24 733
 FT DOMAIN 24 733 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 734 754
 FT DOMAIN 755 861 POTENTIAL.
 FT DOMAIN 50 482 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 502 551 SEMA.
 FT DOMAIN 555 636 PSI.
 FT DISULFID 576 624 IG-LIKE C2-TYPE.
 FT CARBOHYD 49 49 BY SIMILARITY.
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 632 632 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 861 AA; 95714 MW; 533CD6D271A6D79B CRC64;

Alignment Scores:
 Pred. No.: 1-6 Length: 861
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4D_MOUSE (1-861)
 QY 914 GAAGATGACAAAGTATATCTTTCTTC 940
 |||||
 Db 227 GluaspaspysValtyrPhePhePhe 235
 |||||

RESULT 26
 SM4D HUMAN
 ID SM4D HUMAN STANDARD; PRT; 862 AA.
 AC Q92854;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4D precursor (Leukocyte activation antigen CD100) (BB18)
 DE (AB) (GR3).
 GN SEMA4D OR CD100.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=97030273; PubMed=8876214;
 RA Hall K.T., Boumsell L., Schultze J.L., Bousiotis V.A., Dorfman D.M.,
 RA Cardoso A.A., Bensussan A., Nadler L.M., Freeman G.J.;
 RA "Human CD100, a novel leukocyte semaphorin that promotes B-cell
 RT aggregation and differentiation.";
 RT Proc. Natl. Acad. Sci. U.S.A. 93:11780-11785(1996).
 CC -!- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL
 CC AS IN THE NERVOUS SYSTEM. INDUCES B CELLS TO AGGREGATE AND
 CC IMPROVES THEIR VIABILITY IN VITRO.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN SKELETAL MUSCLE,
 CC PERIPHERAL BLOOD LYMPHOCYTES, SPLEEN, AND THYMUS AND ALSO
 CC EXPRESSED AT LOWER LEVELS IN TESTES, BRAIN, KIDNEY, SMALL
 CC INTESTINE, PROSTATE, HEART, PLACENTA, LUNG, AND PANCREAS BUT NOT
 CC IN COLON OR LIVER.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD100 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd100.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U60800; AAC50810.1; -;
 CC Genew; HGNC:10732; SEMA4D.
 CC MIM; 601866; -;
 CC GO; GO:0006916; P:anti-apoptosis; TAS.
 CC GO; GO:0007155; P:cell adhesion; TAS.
 CC GO; GO:0006955; P:immune response; TAS.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; plexin-like.
 CC InterPro; IPR002165; plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS00835; IG-LIKE; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 21
 FT CHAIN 22 862
 FT DOMAIN 22 734
 FT TRANSMEM 735 755
 FT DOMAIN 756 862
 FT DOMAIN 50 482
 FT DOMAIN 502 551
 FT DOMAIN 554 636
 FT DISULFID 576 624
 FT CARBOHYD 49 49
 FT CARBOHYD 77 77
 FT CARBOHYD 139 139
 FT CARBOHYD 191 191
 FT CARBOHYD 329 329
 FT CARBOHYD 379 379
 FT CARBOHYD 419 419
 FT CARBOHYD 613 613
 FT CARBOHYD 632 632
 SQ SEQUENCE 862 AA; 96149 MW; 7B18EFA98789371 CRC64;

Alignment Scores:
 Pred. No.: 1.6 Length: 862
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4D_HUMAN (1-862)

QY 1055 TTCTCAAGCTCGTCTGATTTCCTCA 1081

Db 274 PheLeuLysAlaArgLeuIleCysSer 282

RESULT 27

IF38 HUMAN

ID IF38 HUMAN STANDARD; PRT; 913 AA.

AC Q99613; Q00215;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Eukaryotic translation initiation factor 3 subunit 8 (eIF3 p110)

DE (eIF3c).

GN EIF38.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
 RP MEDLINE=97150873; PubMed=8995409;
 RX Asano K., Kinzy T.G., Merrick W.C., Hershey J.W.B.;
 RA "Conservation and diversity of eukaryotic translation initiation
 RT factor eIF3.";
 RL J. Biol. Chem. 272:1101-1109(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
 METHIONYL-TRNAI AND MRNA.
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 12 DIFFERENT SUBUNITS.
 CC -1- DOMAIN: CONTAINS 1 PCI DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE EIF38 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; U46025; AAD03462.1; -;
 EMBL; AC002544; AAC27426.1; -;
 EMBL; U91326; AAC27674.1; -;
 EMBL; AK000739; BAA91352.1; -;
 EMBL; BC001571; AAH01571.1; -;
 Genew; HGNC:3279; EIF38.

GK; Q99613; -;

MIM; 603916; -;

GO; GO:0005852; C:eukaryotic translation initiation factor 3 . . . ; TAS.

GO; GO:0003743; F:translation initiation factor activity; TAS.

GO; GO:0006446; P:regulation of translational initiation; TAS.

DR InterPro; IPR000717; PCI.
DR Pfam; PF03399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor; Protein biosynthesis.
FT DOMAIN 164 189 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 243 246 POLY-GLU.
FT DOMAIN 291 294 POLY-GLU.
FT CONFLICT 313 314 EK -> VR (IN REF. 2).
SQ SEQUENCE 913 AA; 105343 MW; CB5029F4EB51C1AA CRC64;

Alignment Scores:
Pred. No.: 1.6 Length: 913
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x IF38_HUMAN (1-913)

QY 2093 CAAGCCCTTCTGCTACGTAGTCTACAA 2119
|||||
DB 648 GlnGlyLeuLeuLeuArgSerLeuGln 656

RESULT 28

SLA1_YEAST STANDARD; PRT; 1244 AA.
AC P32790;

DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoskeleton assembly control protein SLA1.
GN SLA1 OR YBL007C OR YBL0321.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DDY 228;
RX MEDLINE=93328765; PubMed=8335689;
RA Holtzman D.A., Yang S., Drubin D.G.;
RT "Synthetic-lethal interactions identify two novel genes, SLA1 and
RT SLA2, that control membrane cytoskeleton assembly in Saccharomycetes
RT cerevisiae."
RL J. Cell Biol. 122:635-644 (1993).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;
RX MEDLINE=93070613; PubMed=1441753;
RA Delaveau T., Jacq C., Perea J.;
RT PDR-like gene and several new open reading frames. a
RL Yeast 8:761-768 (1992).

CC -!- FUNCTION: ESSENTIAL FOR THE PROPER FORMATION OF THE CORTICAL
CC ACTIN CYTOSKELETON. INVOLVED IN CONTROLLING THE SIZE OF CORTICAL
CC PATCHES PERHAPS BY REGULATING THE NUCLEATION OF FILAMENTS AT THE
CC CORTEX.

CC -!- SIMILARITY: SOME TO SEA URCHIN BINDIN.

CC -!- SIMILARITY: Contains 3 SH3 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; 222810; CAA80463.1; -;
DR EMBL; 235768; CAA84826.1; -;
DR EMBL; 547695; AAB23985.1; -;
DR PIR; S25327; S25327.

DR HSSP; Q03526; LAWJ.
DR SGD; S0000103; SLA1.
DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.
DR GO; GO:0007015; P:actin filament organization; IMP.
DR GO; GO:0007047; P:cell wall organization and biogenesis; IMP.
DR GO; GO:0006897; P:endocytosis; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 3.
DR Pfam; PF03983; SHD1; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 3.
DR SMART; SM00326; SH3; 3.
DR PROSITE; PS00002; SH3; 3.
KW Cytoskeleton; Actin-binding; SH3 domain; Repeat.
FT DOMAIN 8 69
FT DOMAIN 70 132
FT DOMAIN 353 415
FT DOMAIN 868 1205
FT REPEAT 868 874
FT REPEAT 877 883
FT REPEAT 887 893
FT REPEAT 923 929
FT REPEAT 945 951
FT REPEAT 1003 1009
FT REPEAT 1020 1026
FT REPEAT 1031 1037
FT REPEAT 1048 1054
FT REPEAT 1065 1071
FT REPEAT 1084 1090
FT REPEAT 1129 1135
FT REPEAT 1155 1161
FT REPEAT 1170 1176
FT REPEAT 1185 1191
FT REPEAT 1200 1206
SQ SEQUENCE 1244 AA; 135848 MW; 7FD85AA776407624 CRC64;

16 X 7 AA APPROXIMATE REPEATS OF
T-G-G-A-M-N-P.

Alignment Scores:

Pred. No.: 1.55 Length: 1244
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SLA1_YEAST (1-1244)

QY 380 CTTTGGATGAGAACGAGTAGCTG 406
|||||

Db 607 LeuLeuAspGluGluArgSerArgLeu 615
|||||

RESULT 29

RYR3_HUMAN

ID RYR3_HUMAN STANDARD; PRT; 4870 AA.

AC Q15413; Q15175; Q15412;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ryanodine receptor 3 (Brain-type ryanodine receptor) (RYR3) (RYR-3)
DE (Brain ryanodine receptor-calcium release channel).

GN RYR3 OR HBRR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, AND VARIANTS.

RC TISSUE=Fetal brain;

RX MEDLINE=98175492; PubMed=9515741;

RA Leeb T., Brenig B.;

RT "cDNA cloning and sequencing of the human ryanodine receptor type 3
RT (RYR3) reveals a novel alternative splice site in the RYR3 gene.";

FT VARIANT 693 693 C -> Y.
FT VARIANT 731 731 /FTID=VAR_011404.
FT VARIANT 1380 1380 V -> I.
FT VARIANT 1380 1380 /FTID=VAR_011405.
FT VARIANT 2268 2268 E -> I.
FT VARIANT 2268 2268 /FTID=VAR_011406.
FT CONFLICT 261 261 MISSING.
FT CONFLICT 494 494 /FTID=VAR_011407.
FT CONFLICT 742 742 S -> R (IN REF. 2).
FT CONFLICT 932 932 I -> V (IN REF. 2).
FT CONFLICT 1081 1081 VSCDLGVPSISPRINGOPVGMF -> GKLLPGRGCPA
FT CONFLICT 1336 1336 SHSASMGSPCRGCL (IN REF. 2).
FT CONFLICT 1480 1480 A -> T (IN REF. 2).
FT CONFLICT 2355 2355 A -> P (IN REF. 2).
FT CONFLICT 2433 2433 K -> E (IN REF. 2).
FT CONFLICT 2546 2433 A -> G (IN REF. 2).
FT CONFLICT 2580 2580 A -> G (IN REF. 2).
FT CONFLICT 2817 2817 F -> I (IN REF. 2).
FT CONFLICT 3684 3684 T -> S (IN REF. 2).
FT CONFLICT 3698 3698 A -> G (IN REF. 2).
FT CONFLICT 4026 4026 S -> P (IN REF. 2).
FT CONFLICT 4083 4083 F -> S (IN REF. 2).
FT CONFLICT 4317 4317 E -> G (IN REF. 2 AND 4).
FT CONFLICT 4537 4537 E -> G (IN REF. 2).
FT CONFLICT 4604 4604 R -> P (IN REF. 2).
FT CONFLICT 4709 4709 I -> L (IN REF. 2).
FT CONFLICT 4870 4870 C -> R (IN REF. 5).
SQ SEQUENCE 4870 AA; 551927 MW; 6D629FB0657390B9 CRC64;

Alignment Scores:

Pred. No.: 1.38 Length: 4870
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RYR3_HUMAN (1-4870)

QY 67 CQTGGTGCACCCCTGACGTGCA 41

Db 2132 ArgGlySerThrProLeuAspValala 2140

RESULT 30

RYR2_HUMAN
ID RYR2_HUMAN STANDARD; PRT; 4967 AA.
AC Q92736; Q15411;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RyR2)
DE (RyR-2) (Cardiac muscle ryanodine receptor-calcium release channel)
DE (hRyR-2).
GN RYR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Heart muscle;
RX MEDLINE=96404895; PubMed=8809036;
RA Tunwell R.E.A., Wickenden C., Bertrand B.M.A., Shevchenko V.I.,
RA Walsh M.B., Allen P.D., Lai F.A.;
RT "The human cardiac muscle ryanodine receptor-calcium release channel:
RT identification, primary structure and topological analysis.";
RL Biochem. J. 318:477-487(1996).
RN [2]
RP SEQUENCE OF 9-87 AND 533-681 FROM N.A., DEVELOPMENTAL STAGE, AND
RP INDUCTION.
RC TISSUE=Heart muscle, and Myometrium;

RX MEDLINE=97220386; PubMed=9148749;
RA Awad S.S., Lamb H.K., Morgan J.M., Dunlop W., Gillespie J.I.;
RT "Differential expression of ryanodine receptor RYR2 mRNA in the non-
RL pregnant and pregnant human myometrium.";
RL Biochem. J. 322:777-783(1997).
RN [3]
RP SEQUENCE OF 4292-4479 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Cerebellum, and Hippocampus;
RX MEDLINE=98268728; PubMed=9607712;
RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;
RT "Partial cloning and differential expression of ryanodine
RT receptor/calcium-release channel genes in human tissues including the
RT hippocampus and cerebellum.";
RL Neuroscience 85:205-216(1998).
RN [4]
RP VARIANTS VTSIP LEU-2246; SER-2474; LYS-4104 AND CYS-4497.
RX PubMed=11208676;
RA Priori S.G., Napolitano C., Tiso N., Memmi M., Vignati G., Bloise R.,
RA Sorrentino V.V., Danieli G.A.;
RT "Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie
RL catecholaminergic polymorphic ventricular tachycardia.";
RL Circulation 103:196-200(2001).
RN [5]
RP VARIANTS VPFT SER-2328; ARG-4201 AND PHE-4653, AND VARIANT ARG-2958.
RX MEDLINE=21112862; PubMed=11157710;
RA Laitinen P.J., Brown K.M., Piippo K., Swan H., Devaney J.M.,
RA Brahmabhatt B., Donarum E.A., Marino M., Tiso N., Viitasalo M.,
RA Toivonen L., Stephan D.A., Kontula K.;
RT "Mutations of the cardiac ryanodine receptor (RyR2) gene in familial
RT polymorphic ventricular tachycardia.";
RL Circulation 103:485-490(2001).
RN [6]
RP VARIANTS ARVD2 PRO-433 AND ILE-2386.
RX MEDLINE=21096894; PubMed=11159936;
RA Tiso N., Stephan D.A., Nava A., Bagattin A., Devaney J.M., Stanchi P.,
RA Thieret G., Brahmabhatt B., Brown K., Baucé B., Muriago M., Basso C.,
RA Thiéne G., Danieli G.A., Rampazzo A.;
RT "Identification of mutations in the cardiac ryanodine receptor gene in
RT families affected with arrhythmic right ventricular
RT cardiomyopathy type 2 (ARVD2).";
RL Hum. Mol. Genet. 10:189-194(2001).
CC -!- FUNCTION: Communication between transverse-tubules and
CC sarcoplasmic reticulum. Contraction of cardiac muscle is triggered
CC by release of calcium ions from SR following depolarization of T-
CC tubules (By similarity).
CC -!- SUBUNIT: Homotetramer (Potential).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q92736-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q92736-2; Sequence=VSP_005953;
CC -!- TISSUE SPECIFICITY: Heart muscle, brain (cerebellum and
CC hippocampus) and placenta.
CC -!- DEVELOPMENTAL STAGE: Expressed in myometrium during pregnancy.
CC -!- INDUCTION: By TGF-beta.
CC -!- DISEASE: Defects in RYR2 are a cause of arrhythmic right
CC ventricular cardiomyopathy type 2 (ARVD2). ARVD2 is an autosomal
CC dominant disease. Characterized by partial degeneration of the
CC myocardium of the right ventricle, electrical instability and
CC sudden death. It is clinically defined by electrocardiographic and
CC angiographic criteria; pathologic findings, replacement of
CC ventricular myocardium with fatty and fibrous elements,
CC preferentially involve the right ventricular free wall.
CC -!- DISEASE: Defects in RYR2 are a cause of stress-induced polymorphic
CC ventricular tachycardia (VTSP) also known as catecholaminergic
CC polymorphic ventricular tachycardia. VTSP is genetic
CC arrhythmic disorder characterized by stress-induced,
CC bidirectional ventricular tachycardia that may degenerate into
CC cardiac arrest and cause sudden death.
CC -!- DISEASE: Defects in RYR2 are a cause of familial polymorphic
CC ventricular tachycardia (VPFT). VPFT is an autosomal-dominant,

inherited disease with a relatively early onset and a mortality rate of approximately 30% by the age of 30 years. Phenotypically, it is characterized by salvos of bidirectional and polymorphic ventricular tachycardias in response to vigorous exercise, with no structural evidence of myocardial disease.

-!- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-!- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.

-!- SIMILARITY: Contains 3 SPRY domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X98330; CAA66975.1; -
 DR EMBL; Y08218; CAA69395.1; -
 DR EMBL; X91869; CAA62975.1; -
 DR EMBL; A0002511; CAA05502.1; -
 DR PIR; S72269; S72269.
 DR Genew; HGNC:10484; RYR2.
 DR MIM; 180902; -
 DR MIM; 192605; -
 DR MIM; 600996; -
 DR MIM; 604772; -
 DR GO; GO:0015278; P:calcium-release channel activity; TAS.
 DR GO; GO:0006936; P:muscle contraction; TAS.
 DR GO; GO:0008016; P:regulation of heart; TAS.
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001882; Ca/Na_Pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00054; Efh; 2.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR Ionic channel; Receptor; Calcium channel; Calmodulin-binding; Repeat; Transmembrane; Glycoprotein; Phosphorylation; Alternative splicing; Polymorphism; Disease mutation; Cardiomyopathy; Multigene family. CYTOPLASMIC (BY SIMILARITY).

DOMAIN 1 3089
 TRANSMEM 3090 3109
 TRANSMEM 3153 3171 M' (POTENTIAL).
 TRANSMEM 3940 3959 M1 (POTENTIAL).
 TRANSMEM 3978 3995 M2 (POTENTIAL).
 TRANSMEM 4233 4256 M3 (POTENTIAL).
 TRANSMEM 4294 4314 M4 (POTENTIAL).
 TRANSMEM 4500 4520 M5 (POTENTIAL).
 TRANSMEM 4578 4600 M6 (POTENTIAL).
 TRANSMEM 4720 4740 M7 (POTENTIAL).
 TRANSMEM 4768 4786 M8 (POTENTIAL).
 TRANSMEM 4810 4827 M9 (POTENTIAL).
 TRANSMEM 4845 4867 M10 (POTENTIAL).

FT DOMAIN 670 808
 FT DOMAIN 1098 1221
 FT DOMAIN 1423 1561
 FT DOMAIN 4414 4455
 FT DOMAIN 853 966
 FT REPEAT 967 1080
 FT REPEAT 2692 2810
 FT REPEAT 2812 2925
 FT BINDING 2818 3015
 FT BINDING 2774 2806
 FT BINDING 2876 2897
 FT BINDING 2997 3015
 FT MOD_RES 2808 2808
 FT CARBOHYD 3427 3427
 FT CARBOHYD 3536 3536
 FT CARBOHYD 3770 3770
 FT CARBOHYD 3864 3864
 FT CARBOHYD 3865 3865
 FT CARBOHYD 3905 3905
 FT CARBOHYD 4075 4075
 FT CARBOHYD 4104 4104
 FT CARBOHYD 4111 4111
 FT CARBOHYD 4794 4794
 FT VARSPLIC 3715 3715
 FT VARIANT 433 433

Alignment Scores:
 Pred. No.: 1.38 Length: 4967
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RYR2_HUMAN (1-4967)
 Qy 67 CGTGGTGGACACCCCTGAGCTGGCA 41.
 Db 2235 ArgGlySerThrProLeuAspValAla 2243
 |||||
 |||||

RESULT 31
 RYR2_RABIT
 ID_RYR2_RABIT STANDARD; PRT; 4969 AA.
 AC P30957;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RYR2)
 DE (RYR-2) (Cardiac muscle ryanodine receptor-calcium release channel).
 GN RYR2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart muscle;
 RX MEDLINE=90337947; PubMed=2380170;
 RA Otsu K., Willard H.F., Khanna V.K., Zorzato F., Green N.M.,
 MacLennan D.H.;
 RT "Molecular cloning of cDNA encoding the Ca2+ release channel
 RT (ryanodine receptor) of rabbit cardiac muscle sarcoplasmic
 RT reticulum.";
 RL J. Biol. Chem. 265:13472-13483 (1990).
 RN [2]
 RP PHOSPHORYLATION OF SER-2809.
 RX MEDLINE=91250425; PubMed=1645727;
 RA Witcher D.R., Kovacs R.J., Schulman H., Cefali D.C., Jones L.R.;
 RT "Unique phosphorylation site on the cardiac ryanodine receptor
 RT regulates calcium channel activity.";

J. Biol. Chem. 266:11144-11152(1991).

-!- FUNCTION: Communication between transverse-tubules and sarcoplasmic reticulum. Contraction of cardiac muscle is triggered by release of calcium ions from SR following depolarization of T-tubules.

-!- SUBUNIT: Homotetramer (Potential).

-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

-!- TISSUE SPECIFICITY: HEART AND BRAIN.

-!- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-!- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M59743; AAA31179.1; -

PIR; A37113; A37113.

InterPro; IPR000899; Ca-rel_channel.

InterPro; IPR001682; Ca/Na_pore.

InterPro; IPR002048; EF-hand.

InterPro; IPR005821; Ion trans.

InterPro; IPR003608; MIR.

InterPro; IPR001215; Ryanodn_receptor.

InterPro; IPR003032; RYR.

InterPro; IPR003877; SPRY_receptor.

Pfam; PF00036; ehand; 2.

Pfam; PF00520; ion trans; 1.

Pfam; PF02815; MIR; 4.

Pfam; PF01365; RYDR_ITPR; 2.

Pfam; PF02026; RYR; 4.

Pfam; PF00622; SPRY; 3.

PRINTS; PR00795; RYANODINER.

SMART; SM00054; Efh; 2.

SMART; SM00472; MIR; 4.

SMART; SM00449; SPRY; 3.

Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat; Phosphorylation; Glycoprotein; Multigene family; Calmodulin-binding.

DOMAIN 1 3090

TRANSMEM 3091 3110

M' (POTENTIAL).

M1 (POTENTIAL).

M2 (POTENTIAL).

M3 (POTENTIAL).

M4 (POTENTIAL).

M5 (POTENTIAL).

M6 (POTENTIAL).

M7 (POTENTIAL).

M8 (POTENTIAL).

M9 (POTENTIAL).

M10 (POTENTIAL).

SPRY 1.

SPRY 2.

SPRY 3.

1.

2.

3.

4.

MODULATOR (POTENTIAL).

CALMODULIN (POTENTIAL).

FT BINDING 2877 2898 CALMODULIN (POTENTIAL).

FT BINDING 2998 3016 CALMODULIN (POTENTIAL).

FT MOD RES 2809 2809 PHOSPHORYLATION (BY CAMK).

FT CARBOHYD 4105 4105 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 4796 4796 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 4969 AA; 565060 MW; FF6E0684B974BB4D CRC64;

Alignment Scores:

Pred. No.: 1.38 Length: 4969

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.03% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RYR2_RABIT (1-4969)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41

Db 2235 ArgGlySerThrProLeuAspValala 2243

RESULT 32

PND1_ECOLI STANDARD; PRT; 50 AA.

ID_PND1_ECOLI

AC P11902;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE PND protein.

GN PND.

OS Escherichia coli.

OG Plasmid R16.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89150247; PubMed=2465777;

RA Sakikawa T., Akimoto S., Ohnishi Y.;

RT "The pnd gene in E. coli plasmid R16: nucleotide sequence and gene expression leading to cell Mg2+ release and stable RNA degradation.";

RL Biochim. Biophys. Acta 1087:158-166(1989).

CC -!- FUNCTION: WHEN OVEREXPRESSED KILL THE CELLS FROM THE INSIDE BY INTERFERING WITH A VITAL FUNCTION IN THE CELL MEMBRANE.

CC -!- FUNCTION: Involved in cell Mg(2+) release and stable RNA degradation.

CC -!- INDUCTION: IN THE PRESENCE OF RIFAMPICIN AT 42 DEGREES CELSIUS.

CC -!- SIMILARITY: BELONGS TO THE HOK/GEF FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X12833; CAA31320.1; -

PIR; S02763; S02763.

InterPro; IPR000021; Hok/gef_toxin.

Pfam; PF01848; HOK_GEF; 1.

PRINTS; PR00281; HOKGEFTOXIC.

PROSITE; PS00556; HOK_GEF; 1.

KW Transmembrane; Plasmid.

SPRY 2.

SQ SEQUENCE 50 AA; 5796 MW; F4757270A3F21792 CRC64;

Alignment Scores:

Pred. No.: 24.2 Length: 50

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.89% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PND1_ECOLI (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
 |||
 3 GlnArgThrPheLeuMetMetLeu 10

RESULT 33

ID	PND2_ECOLI	STANDARD	PRT	50 AA
AC	P16477			
DT	01-AUG-1990	(Rel. 15, Created)		
DT	01-AUG-1990	(Rel. 15, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	PND protein			
GN	PND2			
OS	Escherichia coli			
OG	Plasmid Incil R483			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales			
OC	Enterobacteriaceae; Escherichia			
OX	NCBI_TaxID=562			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88174457; PubMed=3328043			
RA	Ono K., Akimoto S., Ohnishi Y.			
RT	"Nucleotide sequence of the pnd gene in plasmid R483 and role of the pnd gene product in plasmolysis."			
RL	Microbiol. Immunol. 31:1071-1083(1987)			
CC	-!- FUNCTION: WHEN OVEREXPRESSED KILL THE CELLS FROM THE INSIDE BY INTERFERING WITH A VITAL FUNCTION IN THE CELL MEMBRANE.			
CC	-!- FUNCTION: Involved in cell Mg(2+) release and stable RNA degradation.			
CC	-!- SIMILARITY: BELONGS TO THE HOK/GEF FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; D00364; BAA00270.1; -			
DR	PIR; J00071; J00071			
DR	InterPro; IPR000021; Hok/gef_toxin.			
DR	Pfam; PF01848; HOK_GEF; 1.			
DR	PRINTS; PR00281; HOKGEFTOXIC.			
DR	PROSITE; PS00556; HOK_GEF; 1.			
KW	Transmembrane; Plasmid.			
FT	TRANSMEM 5 25			
SQ	SEQUENCE 50 AA; 5775 MW; 2DA904DCBE81AFA9 CRC64;			

Alignment Scores:

Pred. No.:	24.2	Length:	50
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	1	Gaps:	0

US-09-774-490-1 (1-2709) x PND2_ECOLI (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
 |||
 3 GlnArgThrPheLeuMetMetLeu 10

RESULT 34

ID	MTRB_BACSU	STANDARD	PRT	75 AA
AC	P19466			
DT	01-FEB-1991	(Rel. 17, Created)		
DT	01-FEB-1991	(Rel. 17, Last sequence update)		

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transcription attenuation protein mtrB (Tryptophan RNA-binding

DE attenuator protein) (Trp RNA-binding attenuation protein) (TRAP).

GN MTRB.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91062353; PubMed=2123343;

RA Gollnick P., Ishino S., Kuroda M.I., Henner D.J., Yanofsky C.;

RT "The mtr locus is a two-gene operon required for transcription

RT attenuation in the trp operon of Bacillus subtilis."

RL Proc. Natl. Acad. Sci. U.S.A. 87:8726-8730(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guisepi G., Guy B.J., Haga K., Haesch J., Harwood C.R., Henaut A.,

RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Seguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tanakoshi A., Tanaka T., Terpatra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis."

RL Nature 390:249-256(1997).

RN [3]

RP FUNCTION.

RX MEDLINE=92202128; PubMed=1551827;

RA Babitzke P., Gollnick P., Yanofsky C.;

RT "The mtrAB operon of Bacillus subtilis encodes GTP cyclohydrolase I

RT (MtrA), an enzyme involved in folic acid biosynthesis, and MtrB, a

RT regulator of tryptophan biosynthesis."

RL J. Bacteriol. 174:2059-2064(1992).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=95231646; PubMed=7715723;

RA Antson A.A., Otridge J., Brzozowski A.M., Dodson E.J., Dodson G.G.,

RA Wilson K.S., Smith T.M., Van M., Kurecki T., Gollnick P.;

RT "The structure of trp RNA-binding attenuation protein."

RL Nature 374:693-700(1995).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).

RX MEDLINE=95055737; PubMed=7525975;

RA Antson A.A., Brzozowski A.M., Dodson E.J., Dauter Z., Wilson K.S.,

RA Kurecki T., Otridge J., Gollnick P.;

RT "11-fold symmetry of the trp RNA-binding attenuation protein (TRAP)

RT from Bacillus subtilis determined by X-ray analysis."

RL J. Mol. Biol. 244:1-5(1994).

CC -!- FUNCTION: REQUIRED FOR TRANSCRIPTION ATTENUATION CONTROL IN THE

CC TRP OPERON. THIS TRANS-ACTING FACTOR SEEMS TO RECOGNIZE A 10 BASES
 CC NUCLEOTIDE SEQUENCE IN THE TRP LEADER TRANSCRIPT CAUSING
 CC TRANSCRIPTION TERMINATION. BINDS THE LEADER RNA ONLY IN PRESENCE
 CC OF L-TRYPTOPHAN.
 CC -|- SUBUNIT: OLIGOMER OF 11 IDENTICAL SUBUNITS ARRANGED IN DOUGHNUT-
 CC LIKE STRUCTURE.
 CC -|- SIMILARITY: WITH REGA FROM PHAGE T4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M37320; AAA22616.1; -
 CC EMBL; M80245; AAA20853.1; -
 CC EMBL; Z99115; CAB14193.1; -
 CC PIR; B38256; B38256.
 CC PDB; 1WAP; 03-JUN-95.
 CC Subtilist; BG10278; mtrB.
 CC InterPro; IPR000824; TrpBP.
 CC Pfam; PF02081; TrpBP; 1.
 CC PRINTS; PR00687; TRPRNAP.
 CC ProDom; PD027918; TrpBP; 1.
 CC Transcription regulation; RNA-binding; 3D-structure;
 CC Complete proteome.
 CC STRAND 9 14
 CC STRAND 19 25
 CC STRAND 32 38
 CC TURN 40 41
 CC STRAND 43 47
 CC STRAND 52 58
 CC STRAND 61 65
 CC TURN 66 67
 CC STRAND 68 72
 CC SEQUENCE 75 AA; 8328 MW; 34C173ABD15A4B31 CRC64;
 KW
 FT STRAND 9 14
 FT STRAND 19 25
 FT STRAND 32 38
 FT TURN 40 41
 FT STRAND 43 47
 FT STRAND 52 58
 FT STRAND 61 65
 FT TURN 66 67
 FT STRAND 68 72
 FT SEQUENCE 75 AA; 8328 MW; 34C173ABD15A4B31 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 23.3 Length: 75
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x MTRB_BACSU (1-75)
 Qy 2034 GCGGAATGAGAGCGAAAGAAG 2057
 Db 68 GlyGluMetLysSerGluLysLys 75
 RESULT 35
 VSPC_TRIGA
 ID VSPC_TRIGA STANDARD; PRT; 257 AA.
 AC O13062;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Venom serine proteinase 2C precursor (EC 3.4.21.-).
 GN TG2C.
 OS Trimeresurus gramineus (Indian green tree viper) (Green habu snake).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;
 CC Viperidae; Crotalinae; Trimeresurus.
 CC NCBI_TaxID=8767;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Venom gland;
 RX MEDLINE=97096898; PubMed=8941719;
 RA Doshinawu M., Ogawa T., Nakashima K., Nobuhisa I., Chijiwa T.,
 RA Shimomigaishi Y., Fukumaki Y., Niwa M., Yamashina I., Hattori S.,

RA Ohno M.;
 RT "Accelerated evolution of crotalinae snake venom gland serine
 RT proteases.";
 RL FEBS Lett. 397:83-88(1996).
 CC -|- FUNCTION: Thrombin-like snake venom serine protease.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. SNAKE VENOM SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D67084; BAA19982.1; -
 CC HSPF; P00763; LDPO.
 CC MEROPS; S01.185; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
 CC SIGNAL 1 18
 CC PROPEP 19 24
 CC CHAIN 25 257
 CC ACT_SITE 64 64
 CC ACT_SITE 109 109
 CC ACT_SITE 203 203
 CC DISULFID 31 162
 CC DISULFID 49 65
 CC DISULFID 97 255
 CC DISULFID 141 209
 CC DISULFID 173 188
 CC DISULFID 199 224
 CC CARBOHYD 116 116
 CC CARBOHYD 120 120
 CC CARBOHYD 121 121
 CC SEQUENCE 257 AA; 28042 MW; 2F50B6947CB33AB1 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 21 Length: 257
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x VSPC_TRIGA (1-257)
 Qy 1998 TTCGACTGCATTCACAAAATGTG 1975
 Db 74 PheGlyLeuHisSerLysAsnVal 81
 RESULT 36
 GRBB_BACSU
 ID GRBB_BACSU STANDARD; PRT; 368 AA.
 AC P39570;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Spore germination protein B2.
 GN GERBB.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC NCBI_TaxID=1423;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=94282292; PubMed=8012571;
 RA Corfe B.M., Sammons R.L., Smith D.A., Maueel C.;
 RT "The gerB region of the Bacillus subtilis 168 chromosome encodes a
 RT homologue of the gerA spore germination operon";
 RL Microbiology 140:471-478(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boursier L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brulliet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Evington J., Fabret C., Ferrich S., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: INVOLVED IN THE RESPONSE TO THE GERMINATIVE MIXTURE OF
 CC L-ASPARAGINE, GLUCOSE, FRUCTOSE AND POTASSIUM IONS (AGFK). COULD
 CC BE AN AMINO ACID TRANSPORTER. CANNOT STIMULATE GERMINATION IN THE
 CC ABSENCE OF GERD AND GERK GENE PRODUCTS (FRUCTOSE AND GLUCOSE
 CC RECEPTORS RESPECTIVELY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE FORESPORE COMPARTMENT OF THE
 CC DEVELOPING SPORANGIUM.
 CC -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
 CC PERMEASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L16960; AAA22467.1; -;
 CC PIR; 299122; CABL5598.1; -;
 CC PIR; I39856; I39856.
 CC Subtilist; BG10641; gerB.
 CC InterPro; IPR002293; AA/rel_permease1.
 CC InterPro; IPR002422; AA/rel_permease2.
 CC InterPro; IPR002091; ArAA_permease.
 CC InterPro; IPR004761; Spore_permease.
 CC Pfam; PF03845; Spore_permease; 1.
 CC PRINTS; PR00186; ARONAPRMEASE.
 CC TIGRFAMs; TIGR00912; 2A0309; 1.
 KW Germination; Transport; Amino-acid transport; Transmembrane;

KW Complete proteome. 30 POTENTIAL.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 43 63 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 145 165 POTENTIAL.
 FT TRANSMEM 187 207 POTENTIAL.
 FT TRANSMEM 217 237 POTENTIAL.
 FT TRANSMEM 247 267 POTENTIAL.
 FT TRANSMEM 282 302 POTENTIAL.
 FT TRANSMEM 308 328 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 SQ SEQUENCE 368 AA; 41709 MW; AB6B769512B16806 CRC64;
 Alignment Scores:
 Pred. No.: 20.3 Length: 368
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x GRBB_BACSU (1-368)
 QY 2481 GTCTCTCTGTTTCTACCTTTCTTA 2458
 DB 56 ValLeuLeuPheLeuProPheLeu 63
 RESULT 37
 ID SPB8 HUMAN STANDARD; PRT; 374 AA.
 AC P50452;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytoplasmic antiprotease 2 (CAP2) (CAP-2) (Protease inhibitor 8)
 DE (Serpin B8).
 GN SERPINB8 OR P18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kisiel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine proteinase
 RT inhibitors";
 RL J. Biol. Chem. 270:29854-29861(1995).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L40377; AAC41939.1; -;
 CC PIR; A59273; A59273.
 CC HSSP; P05120; 1BY7.
 CC Genew; HGNC:8952; SERPINB8.
 CC MIM; 601697; -;
 CC GO; GO:0005829; Cytosol; TAS.
 CC GO; GO:0005515; F:protein binding activity; TAS.
 CC GO; GO:0004868; F:serpin; TAS.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.

```

DR SMART, SMO0093; SERPIN; 1.
KW PROSITE; PS00284; SERPIN; 1.
FT ACT SITE 339 340 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;

Alignment Scores:
Pred. No.: 20.3 Length: 374
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SPB8_HUMAN (1-374)
QY 1590 TTCCTAAGGACACTGGTATGATT 1613
Db 282 PheLeuArgA-gLeuGlyMetile 289

RESULT 38
SH1D_CANFA STANDARD; PRT; 377 AA.
AC P11614;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 5-hydroxytryptamine 1D receptor (5-HT-1D) (Serotonin receptor).
GN HTR1D OR RDC4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=89242119; PubMed=2541503;
RA Libert F., Parmentier M., Lefort A., Dumont J.E., Vassart G.;
RA Maenhaut C., Simons M.-J., Dumont J.E., Vassart G.;
RT "Selective amplification and cloning of four new members of the G
protein-coupled receptor family.";
RL Science 244:569-572(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=90245610; PubMed=2159630;
RA Libert F., Parmentier M., Lefort A., Dumont J.E., Vassart G.;
RT "Complete nucleotide sequence of a putative G protein coupled
receptor: RDC4.";
RL Nucleic Acids Res. 18:1916-1916(1990).
RN [3]
RP FUNCTION.
RX MEDLINE=92062181; PubMed=1659418;
RA Maenhaut C., van Sande J., Maesart C., Dinsart C., Libert F.,
RA Monferini E., Giraldo E., Ladinsky H., Vassart G., Dumont J.E.;
RT "The orphan receptor cDNA RDC4 encodes a 5-HT1D serotonin receptor.";
RL Biochem. Biophys. Res. Commun. 180:1460-1468(1991).
RN [4]
RP FUNCTION.
RX MEDLINE=92100052; PubMed=1758439;
RA Zgombick J.M., Weinshank R.L., Macchi M., Schechter L.E.,
RA Branchek T.A., Hartig P.R.;
RT "Expression and pharmacological characterization of a canine 5-
hydroxytryptamine1D receptor subtype.";
RL Mol. Pharmacol. 40:1036-1042(1991).
CC -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CYCLASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.

```

----- THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch). -----

EMBL; X14049; CAA32207.1; --
PIR; B30341; B30341.
InterPro; IPR000276; GPCR_Rhodpsn.
Pfam; PF00001; 7tm1; 1.
PRINTS; PR00237; GPCRHOPOFSN.
PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Multigene family.
FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 62 1 (POTENTIAL).
FT DOMAIN 63 75 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 76 98 2 (POTENTIAL).
FT DOMAIN 99 108 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 109 134 3 (POTENTIAL).
FT DOMAIN 135 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 176 4 (POTENTIAL).
FT DOMAIN 177 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 218 5 (POTENTIAL).
FT DOMAIN 219 302 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 303 326 6 (POTENTIAL).
FT DOMAIN 327 335 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 336 360 7 (POTENTIAL).
FT DOMAIN 361 377 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 111 188 BY SIMILARITY.
SQ SEQUENCE 377 AA; 41882 MW; 8B6406DCE2123EE4 CRC64;

Alignment Scores:
Pred. No.: 20.3 Length: 377
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SH1D_CANFA (1-377)
QY 1917 ATAATCACCATTGCGCACAGCCCTG 1940
Db 47 IleIleThrMetAlaThrAlaLeu 54

RESULT 39
ID _BMP_SOCMV STANDARD; PRT; 462 AA.
AC P15628;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Inclusion body matrix protein (Viroplasma)
GN VI.
OS Soybean chlorotic mottle virus.
OC Viruses; Retroviruses; Caulimoviridae;
OC Soybean chlorotic mottle-like viruses.
OX NCBI_TaxID=10651;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098857; PubMed=2602148;
RA Hasegawa A., Verver J., Shimada A., Saito M., Goldbach R.,
RA van Kammen A., Miki K., Kameya-Iwaki M., Hibi T.;
RT "The complete sequence of soybean chlorotic mottle virus DNA and the
identification of a novel promoter.";

CC 2 ferrocyclochrome c.
 CC -1- COFACTOR: CONTAINS TWO FAD AND FOUR TO SIX ZINC MOLES PER MOLE.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- FAMILY: BELONGS TO THE FAD-BINDING OXIDOREDUCTASE/TRANSFERASE
 CC FAMILY 4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X71628; CAA50635.1; -
 CC PIR; S51528; S51528.
 CC InterPro; IPR004113; FAD-oxidase C.
 CC InterPro; IPR006094; Oxid FAD bind.
 CC Pfam; PF02913; FAD-oxidase C; 1.
 CC Pfam; PF01565; FAD_binding_4; 1.
 CC Oxidoreductase; Flavoprotein; FAD; Transit peptide; Mitochondrion;
 CC Zinc.
 CC -----
 CC TRANSIT 1 ? MITOCHONDRION.
 CC CHAIN ? 579 D-LACTATE DEHYDROGENASE [CYTOCHROME].
 CC SEQUENCE 579 AA; 63484 MW; 0DE3A07DC4934883 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 19.6 Length: 579
 CC Score: 8.00 Matches: 8
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.89% Indels: 0
 CC DB: 1 Gaps: 0
 CC -----
 CC US-09-774-490-1 (1-2709) x DLD1_KLJULA (1-579)
 CC -----
 CC QY 495 GAGAGATCAATCGACAGTGGCTG 518
 CC DB 470 GluGluMetAsnAlaSerGlyLeu 477
 CC -----
 CC RESULT 42
 CC ID_YE9C SCHPO STANDARD; PRT; 697 AA.
 CC AC 013773;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Hypothetical J-domain protein C17A5.12 in chromosome 1.
 CC GN SPAC17A5.12.
 CC OS Schizosaccharomyces pombe (Fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC OC Schizosaccharomycetes.
 CC OX NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21848401; PubMed=11859360;
 CC RA Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 CC RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 CC RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 CC RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 CC RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 CC RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 CC RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 CC RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 CC RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E.,
 CC RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 CC RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 CC RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC -1- CAUTION: THE INITIATOR METHIONINE MAY BE FURTHER DOWNSTREAM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z98849; CAB11512.1; -
 CC PIR; T37827; T37827.
 CC GenesDB Spombe; SPAC17A5.12; -
 CC InterPro; IPR001623; DnaJ_N.
 CC InterPro; IPR001440; TPR.
 CC InterPro; IPR000449; UBA_domain.
 CC Pfam; PF00627; UBA; 1.
 CC SMART; SM00271; DnaJ; 1.
 CC PROSITE; PS00076; DnaJ_2; 1.
 CC PROSITE; PS00030; UBA; 1.
 CC KW Hypothetical protein; Chaperone.
 CC FT DOMAIN 181 221 UBA.
 CC FT DOMAIN 633 696 J-DOMAIN.
 CC SQ SEQUENCE 697 AA; 79483 MW; 43789C80AFC18767 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 19.2 Length: 697
 CC Score: 8.00 Matches: 8
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.89% Indels: 0
 CC DB: 1 Gaps: 0
 CC -----
 CC US-09-774-490-1 (1-2709) x YE9C_SCHPO (1-697)
 CC -----
 CC QY 383 TTGATGAGGACGAGTAGGCTG 406
 CC DB 592 LeuAspGluGluArgSerArgLeu 599
 CC -----
 CC RESULT 43
 CC ID_SM4A MOUSE STANDARD; PRT; 760 AA.
 CC AC Q62178;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Semaphorin 4A precursor (Semaphorin B) (Sema B).
 CC GN SEMA4A OR SEMA4 OR SEMB.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=NMRI; TISSUE=Brain;
 CC RX MEDLINE=95267431; PubMed=7748561;
 CC RA Puschel A.W., Adams R.H., Betz H.;
 CC "Marine semaphorin D/collapsin is a member of a diverse gene family
 CC and creates domains inhibitory for axonal extension.";
 CC RL Neuron 14:941-948(1995).
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO

CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
 CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
 CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X85991; CAA59983.1; -;
 CC PIR; I48745; I48745.
 CC MGI; 107560; Sema4a.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 KW SIGNAL 1 32
 FT CHAIN 33 760
 FT DOMAIN 33 682
 FT TRANSMEM 683 703
 FT DOMAIN 704 760
 FT DOMAIN 64 478
 FT DOMAIN 496 547
 FT DOMAIN 572 630
 FT DOMAIN 609 612
 FT DOMAIN 579 623
 FT DISULFID 579 623
 FT CARBOHYD 120 120
 FT CARBOHYD 135 135
 FT CARBOHYD 496 496
 FT CARBOHYD 606 606
 FT CARBOHYD 606 606
 SQ SEQUENCE 760 AA; 83458 MW; 0A4D6B80767B910F CRC64;
 Alignment Scores:
 Pred. No.: 19.1 Length: 760
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 Db: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SM4A_MOUSE (1-760)
 QY 578 ACTCACTGTAGCGCTGGGACG 601
 Db 137 ThrHisLeuTyraLacCysGlyThr 144
 |||||
 RESULT 44
 SM4A_HUMAN
 ID SM4A_HUMAN STANDARD; PRT; 761 AA.
 AC Q9H3E1; Q9HUA9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4A precursor (Semaphorin B) (Sema B).
 GN SEMA4A OR SEMB.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.;
 RT "Human semaphorin B.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB029394; BAB20087.1; -;
 CC EMBL; BC020974; AAH20974.1; -;
 CC MIM; 607292; -;
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 KW SIGNAL 1 32
 FT CHAIN 33 761
 FT DOMAIN 33 683
 FT TRANSMEM 684 704
 FT DOMAIN 705 761
 FT DOMAIN 64 478
 FT DOMAIN 496 548
 FT DOMAIN 573 631
 FT DOMAIN 580 624
 FT DISULFID 580 624
 FT CARBOHYD 120 120
 FT CARBOHYD 135 135
 FT CARBOHYD 496 496
 FT CARBOHYD 607 607
 FT CARBOHYD 293 328
 FT CONFLICT 293 354
 FT CONFLICT 354 354
 FT REF. 1)
 FT Y -> F (IN REF. 1).

SQ SEQUENCE 761 AA; 83573 MW; 5BCB889AA32A2BB3 CRC64;

Alignment Scores:

Pred. No.: 19.1 Length: 761
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4A_HUMAN (1-761)

QY 1043 AAATGACACATTCCTCAAGCT 1066

DB 282 LysTrpThrPheLeuLysAla 289

RESULT 45

SM4G_MOUSE

ID SM4G_MOUSE STANDARD; PRT; 837 AA.

AC Q9WHU7;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 4G precursor.

CN SEMA4G.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RA MEDLINE=99425180; PubMed=10495281;

RL Li H., Wu D.K., Sullivan S.L.;

RT "Characterization and expression of sema4g, a novel member of the

semaphorin gene family";

RL Mech. Dev. 87:169-173(1999).

CC -!- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, AND SEVERAL SENSORY ORGANS

CC AS WELL AS SPECIFIC POPULATIONS OF PROJECTION NEURONS.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF134918; AAD30541.1; --

DR MGD; MGI:1347047; Sema4g.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003006; Ig MHC.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR002165; Plexin_repeat.

DR InterPro; IPR001627; Sema.

DR Pfam; PF01437; PSI; 1.

DR Pfam; PF01403; Sema; 1.

DR SMART; SM00409; IG; 1.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; Sema; 1.

DR PROSITE; PS50835; IG_LIKE; 1.

KW Neurogenesis; Developmental protein; Glycoprotein.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 837 SEMAPHORIN 4G.

FT DOMAIN 18 673 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 674 694
 FT DOMAIN 695 837
 FT DOMAIN 56 487
 FT DOMAIN 505 556
 FT DOMAIN 565 647
 FT DOMAIN 563 566
 FT DOMAIN 762 773
 FT DISULFID 582 630
 FT CARBOHYD 55 55
 FT CARBOHYD 111 111
 FT CARBOHYD 126 126
 FT CARBOHYD 286 286
 FT CARBOHYD 540 540
 FT CARBOHYD 596 596
 SQ SEQUENCE 837 AA; 92378 MW; 5C6E9C9CA545EB1 CRC64;

Alignment Scores:

Pred. No.: 18.9 Length: 837
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4G_MOUSE (1-837)

QY 1055 TTCCTCAAGCTCGTCTGATTTC 1078

DB 285 PheLeuLysAlaArgLeuLysCys 292

RESULT 46

SM4G_HUMAN

ID SM4G_HUMAN STANDARD; PRT; 838 AA.

AC Q9NTN9; Q9HCF3;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 4G precursor.

CN SEMA4G OR KIAA1619.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Heath P.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20450683; PubMed=10997877;

RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes.

XVII. The complete sequences of 100 new cDNA clones from brain which

code for large proteins in vitro.";

RL DNA Res. 7:273-281(2000).

CC -!- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL133215; -- NOT ANNOTATED CDS.

DR EMBL; AB046839; BAB13445.1; ALT_INIT.

DR Genew; HGNC:10735; SEMA4G.

DR InterPro; IPR007110; Ig-like.

```

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 838
FT DOMAIN 18 675
FT TRANSMEM 676 696
FT DOMAIN 697 838
FT DOMAIN 56 489
FT DOMAIN 507 558
FT DOMAIN 567 649
FT DOMAIN 565 568
FT DOMAIN 763 774
FT DISULFID 584 632
FT CARBOHYD 55 55
FT CARBOHYD 111 111
FT CARBOHYD 126 126
FT CARBOHYD 388 388
FT CARBOHYD 542 542
FT CARBOHYD 598 598
FT CONFLICT 543 543
SQ SEQUENCE 838 AA; 91496 MW; 9B281AE8681E245 CRC64;

Alignment Scores:
Pred. No.: 18.9 Length: 838
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4G_HUMAN (1-838)
QY 578 ACTCACTTGACGCTGTGGAAC 601
Db 128 ThrHisLeuTyrAlaCysGlyThr 135

RESULT 47
FIBI_PETMA
ID FIBI_PETMA STANDARD; PRT; 966 AA.
AC P02674;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha-1 chain precursor [Contains: Fibrinopeptide A]
DE (Fragment).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90122867; PubMed=2611265;
RA Wang Y.Z., Patterson J., Gray J.E., Yu C., Cottrell B.A., Shimizu A.,
RA Graham D., Riley M., Doolittle R.F.;
RT "Complete sequence of the lamprey fibrinogen alpha chain.";
RL Biochemistry 28:9801-9806(1989).
RN [2]
RP SEQUENCE OF 6-11.
RX MEDLINE=77065679; PubMed=998989;
RA Cottrell B.A., Doolittle R.F.;
RT "Amino acid sequences of lamprey fibrinopeptides A and B and
characterizations of the junctions split by lamprey and mammalian

```

```

thrombins.";
RL Biochim. Biophys. Acta 453:426-438(1976).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: NOT GLYCOSYLATED.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC CONVERTED INTO THE HARD CLOT BY FACTOR XIIIa WHICH CATALYZES THE
CC EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC MONOMERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; M30123; AAA49263.1; -.
CC PIR; A33626; A33626.
CC PDB; 1LWU; 23-AUG-02.
CC PDB; 1N73; 07-JAN-03.
CC Blood coagulation; Plasma; Signal; Repeat; 3D-structure.
KW NON TER 1 1
FT SIGNAL <1 5
FT PEPTIDE 6 11
FT CHAIN 12 966
FT SITE 11 12
FT
FT FIBRINOPEPTIDE A (BY SIMILARITY).
FT FIBRINOGEN ALPHA-1 CHAIN.
FT CLEAVAGE (BY THROMBIN; RELEASE
FT FIBRINOPEPTIDE A) (BY SIMILARITY).
FT 22 X 18 AA APPROXIMATE TANDEM REPEATS.
FT SER-RICH.
FT SER-RICH.
FT SER-RICH.
FT DISULFID 26 26
FT INTERCHAIN (WITH ALPHA CHAIN)
FT (BY SIMILARITY).
FT INTERCHAIN (WITH BETA CHAIN)
FT (BY SIMILARITY).
FT INTERCHAIN (WITH GAMMA CHAIN)
FT (BY SIMILARITY).
FT INTERCHAIN (WITH BETA CHAIN)
FT (BY SIMILARITY).
FT INTERCHAIN (WITH GAMMA CHAIN)
FT (BY SIMILARITY).
FT INTERCHAIN (WITH BETA CHAIN)
FT (BY SIMILARITY).
SQ SEQUENCE 966 AA; 97314 MW; 410520898AA799EE CRC64;

Alignment Scores:
Pred. No.: 18.7 Length: 966
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x FIBI_PETMA (1-966)
QY 153 ACAGGAAGGAAGACTAAAGCAGCA 176
Db 949 ThrGlyArgLysThrLysAlaAla 956

RESULT 48
SM5A_HUMAN
ID SM5A_HUMAN STANDARD; PRT; 1074 AA.
AC Q13591; O60408;
DT 30-MAY-2000 (Rel. 39, Created)

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
 GN SEMA5A OR SEMAF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98125554; PubMed=9464278;
 RA Simmons A.D., Puschel A.W., McPherson J.D., Overhauser J., Lovett M.;
 RT "Molecular cloning and mapping of human semaphorin F from the Cri-du-
 chat candidate interval.";
 RL Biochem. Biophys. Res. Commun. 242:685-691(1998).
 RN [2]
 RP SEQUENCE OF 1-494 FROM N.A.
 RA Kalicki J., Harmon G.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBSJ databases.
 CC -!- FUNCTION: MAY ACT AS POSITIVE AXONAL GUIDANCE CUES.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U52840; AAC09473.1; -.
 DR EMBL; AC004615; AAC14668.1; -.
 DR PIR; JC5928; JC5928.
 DR Genew; HGNC:10736; SEMA5A.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007399; P:neurogenesis; TAS.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp.1; 6.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 6.
 DR PROSITE; PS50092; TSP1; 6.
 DR Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 22
 FT CHAIN 1 22
 FT DOMAIN 23 1074
 FT DOMAIN 23 968
 FT TRANSMEM 969 989
 FT DOMAIN 990 1074
 FT DOMAIN 226 507
 FT DOMAIN 540 593
 FT DOMAIN 595 651
 FT DOMAIN 653 702
 FT DOMAIN 707 765
 FT DOMAIN 784 839
 FT DOMAIN 841 896
 FT DOMAIN 897 944
 FT DOMAIN 142 142
 FT CARBOHYD 168 168
 FT CARBOHYD 227 227
 FT CARBOHYD 277 277
 FT CARBOHYD 323 323
 FT CARBOHYD 367 367
 FT CARBOHYD 437 437

FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 56 56 A -> V (IN REF. 2).
 FT CONFLICT 149 149 A -> T (IN REF. 2).
 FT CONFLICT 382 382 V -> M (IN REF. 2).
 FT CONFLICT 494 494 S -> R (IN REF. 2).
 SQ SEQUENCE 1074 AA; 120570 MW; EE3DB763CBE29407 CRC64;
 Alignment Scores:
 Pred. No.: 18.5 Length: 1074
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SEMA5A_HUMAN (1-1074)
 QY 929 TACTTTTCTTCCTGAAATGCA 952
 DB 230 TyrPhePhePheArgGluAsnAla 237
 RESULT 49
 SEMA5A_MOUSE
 ID SEMA5A_MOUSE STANDARD; PRT; 1077 AA.
 AC Q62217;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
 GN SEMA5A OR SEMAF OR SEMF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI;
 RX MEDLINE=96414430; PubMed=8817451;
 RA Adams R.H., Betz H., Puschel A.W.;
 RT "A novel class of murine semaphorins with homology to thrombospondin
 RT is differentially expressed during early embryogenesis.";
 RL Mech. Dev. 57:33-45(1996).
 CC -!- FUNCTION: MAY ACT AS POSITIVE AXONAL GUIDANCE CUES.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: IN ADULT, DETECTED IN LIVER, BRAIN, KIDNEY,
 CC HEART, LUNG AND SPLEEN.
 CC -!- DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED IN EMBRYONIC AND
 CC ADULT TISSUES. ITS ABUNDANCE DECREASES FROM E10 TO BIRTH..
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X97817; CAA66397.1; -.
 DR MGD; MGI:107556; Sema5a.
 DR GO; GO:0016021; C:integral to membrane; IDA.
 DR GO; GO:0008046; F:axon guidance receptor activity; IDA.
 DR GO; GO:0007411; P:axon guidance; IMP.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01437; PSI; 1.

DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp.1; 5.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 6.
 DR PROSITE; PS0092; TSP1; 6.
 KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 21
 FT CHAIN 22 1077
 FT DOMAIN 22 971
 FT TRANSMEM 972 992
 FT DOMAIN 993 1077
 FT DOMAIN 226 507
 FT DOMAIN 540 593
 FT DOMAIN 595 651
 FT DOMAIN 653 702
 FT DOMAIN 707 765
 FT DOMAIN 784 839
 FT DOMAIN 841 896
 FT DOMAIN 897 944
 FT CARBOHYD 147 147
 FT CARBOHYD 168 168
 FT CARBOHYD 227 227
 FT CARBOHYD 277 277
 FT CARBOHYD 323 323
 FT CARBOHYD 367 367
 FT CARBOHYD 536 536
 FT CARBOHYD 591 591
 FT CARBOHYD 717 717
 FT CARBOHYD 933 933
 SQ SEQUENCE 1077 AA; 120826 MW; 120826 MW; EDABDDDA42789FF CRC64;

Alignment Scores:
 Pred. No.: 19.5 Length: 1077
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SMSA_MOUSE (1-1077)
 QY 929 TACTTTTCTTCGCGAAATGCA 952
 Db 230 TyrPhePheArgGluAsnAla 237
 |||||

RESULT 50
 RPOC_PSEPU
 ID_RPOC_PSEPU STANDARD; PRT; 1409 AA.
 AC P19176;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
 DE beta' chain) (RNA polymerase beta' subunit).
 CN RPOC.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Danilovich A.V., Borodin A.M., Allikmets R.L., Rostapshov V.M.,
 Chernov I.P., Azhikina T.L., Monastyrskaya G.S., Sverdlov E.D.;
 RA "Nucleotide sequence of the rpoC gene coding for the beta'-subunit of
 RT RNA polymerase in Pseudomonas putida";
 RL Dokl. Biochem. 303:241-245 (1998).
 RN [2]
 RP SEQUENCE OF 1-497 FROM N.A.
 RX MEDLINE=89117617; PubMed=3219133;
 RA Borodin A.M., Danilovich A.V., Chernov I.P., Azhikina T.L.,
 Rostapshov V.M., Monastyrskaya G.S.;

RT "Genes coding for RNA polymerase in bacteria. III. The use of
 modified Sanger's method for sequencing the C-terminal region of rpoB
 gene, N-terminal region of rpoC gene and intercistron region of rpoA
 polymerase in Pseudomonas putida.";
 RL Bioorg. Khim. 14:1179-1182 (1988).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA} (N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; X16538; CAA34538.1; -;
 DR EMBL; M38319; AAA25987.1; -;
 DR PIR; JN0420; JN0420.
 DR HSP; Q9KWU6; IHQM.
 DR InterPro; IPR000722; RNA_pol_A.
 DR InterPro; IPR007080; RNA_pol_Rpb1_1.
 DR InterPro; IPR007066; RNA_pol_Rpb1_3.
 DR InterPro; IPR007083; RNA_pol_Rpb1_4.
 DR InterPro; IPR007081; RNA_pol_Rpb1_5.
 DR InterPro; IPR006592; RNA_pol_N.
 DR Pfam; PF04997; RNA_pol_Rpb1_1;
 DR Pfam; PF00623; RNA_pol_Rpb1_2;
 DR Pfam; PF04983; RNA_pol_Rpb1_3;
 DR Pfam; PF05000; RNA_pol_Rpb1_4;
 DR Pfam; PF04998; RNA_pol_Rpb1_5;
 DR SMART; SM00663; RPOLA_N; 1.
 KW Transferase; DNA-directed RNA polymerase; Transcription.
 FT CONFLICT 2 2 P -> L (IN REF. 2).
 FT CONFLICT 17 17 N -> I (IN REF. 2).
 FT CONFLICT 89 89 L -> V (IN REF. 2).
 FT CONFLICT 203 203 I -> T (IN REF. 2).
 FT CONFLICT 328 328 I -> T (IN REF. 2).
 FT CONFLICT 335 335 L -> S (IN REF. 2).
 FT CONFLICT 348 348 L -> R (IN REF. 2).
 FT CONFLICT 482 482 P -> L (IN REF. 2).
 FT CONFLICT 484 484 T -> I (IN REF. 2).
 SQ SEQUENCE 1409 AA; 154544 MW; C5662109D3BE3D81 CRC64;

Alignment Scores:
 Pred. No.: 18.1 Length: 1409
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RPOC_PSEPU (1-1409)
 QY 2527 TCTAGAACCTCAACACAGTAGAA 2550
 Db 971 SerArgAsnLeuLysGlnValGlu 978
 |||||

RESULT 51
 RYR1_PIG
 ID_RYR1_PIG STANDARD; PRT; 5035 AA.
 AC P16960;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ryanodine receptor 1 (skeletal muscle-type ryanodine receptor) (RYR1)

OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=89262082; PubMed=2725677;
 RA Takeshima H., Nishimura S., Matsumoto T., Ishido H., Kangawa K.,
 RA Minamino N., Matsuo H., Ueda M., Hanaoka M., Hirose T., Numa S.;
 RT "Primary structure and expression from complementary DNA of skeletal
 RL muscle ryanodine receptor.";
 RL Nature 339:439-445 (1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=90130482; PubMed=2298749;
 RA Zorzato F., Fujii J., Otsu K., Phillips M., Green N.M., Lai F.A.,
 RA Meissner G., MacLennan D.H.;
 RT "Molecular cloning of cDNA encoding human and rabbit forms of the
 RT Ca2+ release channel (ryanodine receptor) of skeletal muscle
 RT sarcoplasmic reticulum.";
 RL J. Biol. Chem. 265:2244-2256 (1990).
 RN [3]
 RP PHOSPHORYLATION OF SER-2843.
 RX MEDLINE=93120178; PubMed=8380342;
 RA Suko J., Maurer-Fogy I., Plank B., Bertel O., Wiskovsky W.,
 RA Hohenegger M., Hellmann G.;
 RT "Phosphorylation of serine 2843 in ryanodine receptor-calcium release
 RT channel of skeletal muscle by cAMP-, cGMP- and CaM-dependent protein
 RT kinase.";
 RL Biochim. Biophys. Acta 1175:193-206 (1993).
 CC -!- FUNCTION: COMMUNICATION BETWEEN TRANSVERSE-TUBULES AND
 CC SARCOPLASMIC RETICULUM. CONTRACTION OF SKELETAL MUSCLE IS
 CC TRIGGERED BY RELEASE OF CALCIUM IONS FROM SR FOLLOWING
 CC DEPOLARIZATION OF T-TUBULES.
 CC -!- SUBUNIT: Homotetramer (Potential).
 CC -!- TISSUE SPECIFICITY: FAST- OR SLOW-TWITCH SKELETAL MUSCLE.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium
 CC ions, magnesium ions, ATP and calmodulin.
 CC -!- MISCELLANEOUS: The calcium release channel activity resides in the
 CC C-terminal region while the remaining part of the protein
 CC constitutes the 'foot' structure spanning the junctional gap
 CC between the SR and the T-tubule. It is possible that the foot
 CC structure interacts with the cytoplasmic region of the
 CC dihydropyridine receptor.
 CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-
 CC release channel in junctional SR and modulates its activity.
 CC -!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15209; CAA33279.1; -;
 CC EMBL; X15750; CAA33762.1; -;
 CC PIR; S04654; B35041.
 CC InterPro; IPR000699; Ca-rel channel.
 CC InterPro; IPR001882; Ca/Na_Pore.
 CC InterPro; IPR002048; EF-hand.
 CC InterPro; IPR005821; Ion trans.
 CC InterPro; IPR003608; MIR.
 CC InterPro; IPR001215; Ryanodn_receptor.
 CC InterPro; IPR003032; RyR.
 CC InterPro; IPR003877; SPRY_receptor.
 CC Pfam; PF00520; Ion_trans; 1.
 CC Pfam; PF02815; MIR; 4.
 CC Pfam; PF01365; RYDR_ITPR; 2.

DR Pfam; PF02026; RyR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 KW Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
 KW Phosphorylation; Glycoprotein.
 FT DOMAIN 1 3123
 FT TRANSMEM 3124 3144
 FT TRANSMEM 3188 3206
 FT TRANSMEM 3385 4004
 FT TRANSMEM 4023 4041
 FT TRANSMEM 4277 4300
 FT TRANSMEM 4342 4362
 FT TRANSMEM 4559 4580
 FT TRANSMEM 4648 4671
 FT TRANSMEM 4789 4809
 FT TRANSMEM 4837 4856
 FT TRANSMEM 4879 4898
 FT TRANSMEM 4914 4937
 FT DOMAIN 659 797
 FT DOMAIN 1085 1208
 FT DOMAIN 1430 1570
 FT DOMAIN 1873 1913
 FT DOMAIN 842 2959
 FT REPEAT 842 955
 FT REPEAT 956 1069
 FT REPEAT 1345 1360
 FT REPEAT 1373 1388
 FT REPEAT 2726 2845
 FT REPEAT 2846 2959
 FT MOD_RES 2843 2843
 FT MOD_RES 3952 3952
 FT MOD_RES 4323 4323
 FT CARBOHYD 3466 3466
 FT CARBOHYD 3909 3909
 FT CARBOHYD 3950 3950
 FT CARBOHYD 4149 4149
 FT CARBOHYD 4464 4864
 FT CONFLICT 2015 2015
 FT CONFLICT 3481 3485
 SQ SEQUENCE 5037 AA; 565238 MW; 4ABD87AA26697070 CRC64;
 Alignment Scores:
 Pred. No.: 16.2 Length: 5037
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x RYR1_RABIT (1-5037)
 Qy 64 GGGTCGACACCCCTCGACGTGGCA 41
 Db 2369 GlySerThrProLeuaspValala 2276
 RESULT 53
 RYR1_HUMAN
 ID RYR1_HUMAN STANDARD; PRT; 5038 AA.
 AC P21817; Q16314; Q16368; Q9NPK1; Q9PIU4;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ryanodine receptor 1 (skeletal muscle-type ryanodine receptor) (RyR1)
 DE (RyR-1) (skeletal muscle calcium release channel).
 GN RYR1 OR RYDR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND PARTIAL SEQUENCE.

RC TISSUE=Skeletal muscle;
RX MEDLINE=90130482; PubMed=2298749;
RA Zorato F., Fujii J., Otsu K., Phillips M.S., Green N.M., Lai F.A.,
RT Meissner G., McLennan D.H.;
RA "Molecular cloning of cDNA encoding human and rabbit forms of the
RT Ca2+ release channel (ryanodine receptor) of skeletal muscle
RT sarcoplasmic reticulum";
RL J. Biol. Chem. 265:2244-2256(1990).
RN [2]
RN REVISIONS TO 2324; 2840 AND 3380, VARIANT MH ARG-248, AND VARIANTS
RC CYS-471; LEU-1787 AND CYS-2060.
RX TISSUE=Muscle;
RA MEDLINE=92372020; PubMed=1354642;
RX Gillard E.F., Otsu K., Fujii J., Duff C.L., de Leon S., Khanna V.K.,
RA Britt B.A., Worton R.G., McLennan D.H.;
RA "Polymorphisms and deduced amino acid substitutions in the coding
RT sequence of the ryanodine receptor (RYR1) gene in individuals with
RT malignant hyperthermia.";
RL Genomics 13:1247-1254(1992).
RN [3]
RN REVISIONS TO 1365-1368, VARIANT CCD HIS-2435, AND ALTERNATIVE
RC SPLICING.
RX TISSUE=Muscle;
RA MEDLINE=94035117; PubMed=8220422;
RA Zhang Y., Chen H.S., Khanna V.K., de Leon S., Phillips M.S.,
RA Schappert K., Britt B.A., Brownell A.K.W., McLennan D.H.;
RA "A mutation in the human ryanodine receptor gene associated with
RT central core disease.";
RL Nat. Genet. 5:46-50(1993).
RN [4]
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96229657; PubMed=8661021;
RA Phillips M.S., Fujii J., Khanna V.K., de Leon S., Yokobata K.,
RA de Jong P.J., McLennan D.H.;
RA "The structural organization of the human skeletal muscle ryanodine
RT receptor (RYR1) gene.";
RL Genomics 34:24-41(1996).
RN [5]
RN SEQUENCE OF 598-722 FROM N.A.
RX TISSUE=Skeletal muscle;
RA MEDLINE=92347887; PubMed=1639409;
RX Otsu K., Phillips M.S., Khanna V.K., de Leon S., McLennan D.H.;
RT "Refinement of diagnostic assays for a probable causal mutation for
RT porcine and human malignant hyperthermia.";
RL Genomics 13:835-837(1992).
RN [6]
RN SEQUENCE OF 3339-5038 FROM N.A. (ISOFORM 2).
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE OF 4696-4974 FROM N.A.
RX TISSUE=Myometrium;
RA MEDLINE=96032536; PubMed=7556644;
RX Lynn S., Morgan J.M., Lamb H.K., Meissner G., Gillespie J.I.;
RA "Isolation and partial cloning of ryanodine-sensitive Ca2+ release
RT channel protein isoforms from human myometrial smooth muscle.";
RL FEBS Lett. 372:6-12(1995).
RN [8]
RN TISSUE SPECIFICITY.
RX MEDLINE=98268728; PubMed=9607712;
RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;
RA "Partial cloning and differential expression of ryanodine
RT receptor/calcium-release channel genes in human tissues including
RT the hippocampus and cerebellum";
RL Neuroscience 85:205-216(1998).
RN [9]

S-NITROSYLATION OF CYS-3635.
 MEDLINE=21457268; PubMed=11562475;
 Sun J., Xin C., Eu J.P., Stamler J.S., Meissner G.;
 "Cysteine-3635 is responsible for skeletal muscle ryanodine receptor
 modulation by NO";
 Proc. Natl. Acad. Sci. U.S.A. 98:11158-11162(2001).
 [10]
 VN VARIANT MH CYS-614.
 MEDLINE=921128959; PubMed=1774074;
 Gillard E.F., Otsu K., Fujii J., Khanna V.K., de Leon S.,
 Dardemazi J., Britt B.A., Duff C.L., Worton R.G., McLennan D.H.;
 "A substitution of cysteine for arginine 614 in the ryanodine
 receptor is potentially causative of human malignant hyperthermia";
 Genomics 11:751-755(1991).
 [11]
 VN VARIANTS CCD CYS-163 AND MET-403.
 MEDLINE=940351118; PubMed=8220423;
 Quane K.A., Healy J.M.S., Keating K.E., Manning B.M., Couch F.J.,
 Paimucci L.M., Doriguzzi C., Fagerlund T.H., Berg K., Ording H.,
 RA Banducci D., Mortier W., Linz U., Muller C.R., McCarthy T.V.;
 "Mutations in the ryanodine receptor gene in central core disease and
 malignant hyperthermia";
 Nat. Genet. 5:51-55(1993).
 [12]
 VN VARIANT MH SER-522.
 MEDLINE=95130087; PubMed=7829078;
 Quane K.A., Keating K.E., Healy J.M.S., Manning B.M.,
 Krivosic-Horber R., Krivosic I., Monnier N., Lunardi J.,
 McCarthy T.V.;
 "Mutation screening of the RYR1 gene in malignant hyperthermia:
 detection of a novel Tyr to Ser mutation in a pedigree with
 associated central cores";
 Genomics 23:236-239(1994).
 [13]
 VN VARIANT MH ARG-341.
 MEDLINE=94282042; PubMed=80123359;
 Quane K.A., Keating K.E., Manning B.M., Healy J.M.S., Monsieurs K.,
 Heffron J.J.A., Lehane M., Heytens L., Krivosic-Horber R., Adnet P.,
 Ellis F.R., Monnier N., Lunardi J., McCarthy T.V.;
 "Detection of a novel common mutation in the ryanodine receptor gene
 in malignant hyperthermia: implications for diagnosis and
 heterogeneity studies";
 Hum. Mol. Genet. 3:471-476(1994).
 [14]
 VN VARIANT MH ARG-2434.
 MEDLINE=95152512; PubMed=78497112;
 Keating K.E., Quane K.A., Manning B.M., Lehane M., Hartung E.,
 Censler K., Orwyler A., Klausnitzer M., Muller C.R., Heffron J.J.A.,
 McCarthy T.V.;
 "Detection of a novel RYR1 mutation in four malignant hyperthermia
 pedigrees";
 Hum. Mol. Genet. 3:1855-1858(1994).
 [15]
 VN VARIANT MH ARG-2434.
 MEDLINE=95187158; PubMed=7881417;
 Phillips M.S., Khanna V.K., de Leon S., Prodis W., Britt B.A.,
 McLennan D.H.;
 "The substitution of Arg for Gly2433 in the human skeletal muscle
 ryanodine receptor is associated with malignant hyperthermia";
 Hum. Mol. Genet. 3:2181-2186(1994).
 [16]
 VN VARIANT MH CYS-614.
 TISSUE=Blood;
 MEDLINE=95271229; PubMed=7751854;
 Moroni I., Gonano E.G., Comi G.P., Tegazzin V., Prella A., Bordini A.,
 RA Bresolin N., Scarlato F.;
 "Ryanodine receptor gene point mutation and malignant hyperthermia
 susceptibility";
 J. Neurol. 242:127-133(1995).
 [17]
 VN VARIANT MH ARG-35.
 MEDLINE=97219028; PubMed=9066328;
 Lynch P.J., Krivosic-Horber R., Reyford H., Monnier N., Quane K.A.,

RA Adnet P., Haudecoeur G., Krivosic I., McCarthy T.V., Lunardi J.;
 RT "Identification of heterozygous and homozygous individuals with the
 RT novel RYR1 mutation Cys35Arg in a large kindred.";
 RL Anesthesiology 86:620-626(1997).
 RN [18]
 RP VARIANT MH LEU-614.
 RX MEDLINE=98051290; PubMed=9389851;
 RA Quane K.A., Ordling H., Keating K.E., Manning B.M., Heine R.,
 RA Bendixen D., Berg K., Krivosic-Horber R., Lehmann-Horn F.,
 RA Fagerlund T.H., McCarthy T.V.;
 RA "Detection of a novel mutation at amino acid position 614 in the
 RT ryanodine receptor in malignant hyperthermia.";
 RL Br. J. Anaesth. 79:332-337(1997).
 RN [19]
 RP VARIANT MH TRP-552.
 RX MEDLINE=97284075; PubMed=9138151;
 RA Keating K.E., Giblin L., Lynch P.J., Quane K.A., Lehane M.,
 RA Heffron J.J.A., McCarthy T.V.;
 RA "Detection of a novel mutation in the ryanodine receptor gene in an
 RT Irish malignant hyperthermia pedigree: correlation of the IVCT
 RT response with the affected and unaffected haplotypes.";
 RL J. Med. Genet. 34:291-296(1997).
 RN [20]
 RP VARIANTS MH CYS-2163; HIS-2163; MET-2168 AND MET-2206.
 RX MEDLINE=98163444; PubMed=9497245;
 RA Manning B.M., Quane K.A., Ordling H., Urwyler A., Tegazzin V.,
 RA Lehane M., O'Halloran J., Hartung E., Giblin L.M., Lynch P.J.,
 RA Vaughan P., Censier K., Bendixen D., Comi G.P., Heytens L.,
 RA Monsieurs K., Fagerlund T.H., Wolz W., Heffron J.J.A., Mueller C.R.,
 RA McCarthy T.V.;
 RA "Identification of novel mutations in the ryanodine-receptor gene
 RT (RYR1) in malignant hyperthermia: genotype-phenotype correlation.";
 RL Am. J. Hum. Genet. 62:599-609(1998).
 RN [21]
 RP VARIANTS MH CYS-2459 AND HIS-2459.
 RX MEDLINE=98111378; PubMed=9450902;
 RA Manning B.M., Quane K.A., Lynch P.J., Urwyler A., Tegazzin V.,
 RA Krivosic-Horber R., Censier K., Comi G.P., Adnet P., Wolz W.,
 RA Lunardi J., Muller C.R., McCarthy T.V.;
 RA "Novel mutations at a Cpg dinucleotide in the ryanodine receptor in
 RT Alignment Scores:
 Pred. No.: 16.2 Length: 5038
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x RYR1_HUMAN (1-5038)
 QY 64 GGGTCGACCCCTCGACGTGCA 41
 Db 2269 GlySerThrProLeuAspValAla 2276
 RESULT 54
 PSJ1_EUGR
 ID PSJ1_EUGR STANDARD; PRT; 37 AA.
 AC P30394;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Photosystem I reaction center subunit IX (PSI-J).
 PSJ.
 OS Euglena gracilis.
 OG Chloroplast.
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI TaxID=3039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=89041581; PubMed=3141903;
 RA Manzara T.; Hallick R.B.;

RT "Nucleotide sequence of the Euglena gracilis chloroplast genes for
 RT serine and proline transfer RNAs and a functional open reading
 RL frame.";
 RL Nucleic Acids Res. 16:9866-9866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=93169691; PubMed=8435857;
 RA Drager R.G., Hallick R.B.;
 RA "A novel Euglena gracilis chloroplast operon encoding four ATP
 RT synthase subunits and two ribosomal proteins contains 17 introns.";
 RL Curr. Genet. 23:271-280(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=93347989; PubMed=8346031;
 RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
 RA Orsat B., Spielmann A., Stutz E.;
 RA "Complete sequence of Euglena gracilis chloroplast DNA.";
 RL Nucleic Acids Res. 21:3537-3544(1993).
 CC 1- FUNCTION: May help in the organization of the psaE and psaF
 CC subunits (By similarity).
 CC 1- SIMILARITY: Belongs to the psaJ family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z11874; CAA77934.1; -;
 DR EMBL; M18672; AAA84231.1; -;
 DR PIR; X70810; -; NOT_ANNOTATED_CDS.
 DR HAMAP; MF 00522; -; 1.
 DR InterPro; IPR002615; PSI_Psaj.
 DR Pfam; PF01701; PSI_Psaj; 1.
 DR ProDom; PD004198; PSI_Psaj; 1.
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
 FT TRANSMEM 7 27
 FT TRANSMEM 7 27
 SQ SEQUENCE 37 AA; 4296 MW; 3071BEAA56307EB1 CRC64;
 Alignment Scores:
 Pred. No.: 292 Length: 37
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x PSJ1_EUGR (1-37)
 QY 725 CTGACGACATCCCTTTTAATA 745
 Db 21 LeuthralaserLeulle 27
 RESULT 55
 RS14_PYRAE
 ID RS14_PYRAE STANDARD; PRT; 54 AA.
 AC O82VWL; P58732;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S14P.
 GN RPS14P OR PAE2097.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE009857; AAL61943.1;
DR InterPro: IPR001209; Ribosomal_S14.
DR Pfam: PF00253; Ribosomal_S14; 1.
DR PROSITE: PS00527; RIBOSOMAL_S14; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 54 AA; 6444 MW; 153B7BB801EDD963 CRC64;

Alignment Scores:
Pred. No.: 283 Length: 54
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RS14_PYRAE (1-54)

Qy 2406 CTTTCGCCACGCTGTTTCGG 2386
Db 36 LeuCysA-gA-gCysA-gPheA-g 42

RESULT 56
YYP2_SYPN2
ID_YYP2_SYPN2 STANDARD; PRT; 63 AA.
AC P32040;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein in pch 3'region (Fragment).
OS Synecococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32049;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92207922; PubMed=1554697;
RA Schlachter W.M., Bryant D.A.;
RT "Molecular characterization of ferredoxin-NADP+ oxidoreductase in
RT sp. PCC 7002 and studies on the gene product.";
RL Biochemistry 31:3092-3102(1992).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M62334; AAA27325.1;
DR InterPro: IPR001789; Response reg.
DR ProDom: PD000039; Response reg; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Hypothetical protein; Sensory transduction.
FT DOMAIN 20 >63 RESPONSE REGULATORY.

FT NON TER 63 63
SQ SEQUENCE 63 AA; 7142 MW; A9F58D9E01180896 CRC64;

Alignment Scores:
Pred. No.: 279 Length: 63
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YPH2_SYPN2 (1-63)

Qy 110 GCTGACATGGGAGAACAGGC 90
Db 48 AlaAepAenGlyArgThrGly 54

RESULT 57
YYP3_AZOCA
ID_YYP3_AZOCA STANDARD; PRT; 68 AA.
AC P33988;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in RPON 3'region (ORF3) (Fragment).
OS Azorhizobium caulinodans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Hyphomicrobiaceae; Azorhizobium.
OX NCBI_TaxID=7;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ORS571;
RA Stigter J., Schneider M., de Bruijn P.J.;
RT "Azorhizobium caulinodans nitrogen fixation (nif/fix) gene
RT regulation: mutagenesis of the nifA -24/-12 promoter element,
RT characterization of a ntrA(rpoN) gene, and derivation of a model.";
RL Mol. Plant Microbe Interact. 6:238-252(1993).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X69959; CAA49583.1;
DR PIR: S33581; S33581.
KW Hypothetical protein.
FT NON TER 68 68
SQ SEQUENCE 68 AA; 7612 MW; 435C7F141E9C54EB CRC64;

Alignment Scores:
Pred. No.: 277 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YYP3_AZOCA (1-68)

Qy 1719 CTGGGGTTCACGCTCCCTT 1739
Db 24 LeuGlyLeuProSerSerLeu 30

RESULT 58
YL13_ARCFU
ID_YL13_ARCFU STANDARD; PRT; 99 AA.
AC O28167;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF2113.
 GN AF2113.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 CC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., Hickey E.K., Peterson J.D.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Kyrpides N.C.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000958; AAB89154.1; -;
 DR PIR; A69514; A69514.
 DR TIGR; AF2113; -;
 KW Hypothetical protein; Coiled coil; Complete proteome.
 FT DOMAIN 43 95 COILED COIL (POTENTIAL).
 SQ SEQUENCE 99 AA; 11831 MW; 3E9967F3F16AB65C CRC64;
 Alignment Scores:
 Pred. No.: 268 Length: 99
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x YL13_ARCFU (1-99)
 QY 2210 GAGCATTTGGAGACTTCTT 2230
 Db 6 GluHisLeuGluLeuLeu 12
 RESULT 59
 KRF1_COLLI
 ID KRF1_COLLI STANDARD; PRT; 100 AA.
 AC Q9PR15;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Feather keratin Cosl-1/Cosl-3/Cosl-2-1 (F-ker).
 OS Columba livia (Domestic pigeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
 OX NCBI_TaxID=8932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Takahashi R.;
 RT "Pigeon feather keratin genes.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A

CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB017901; BAA33467.1; -;
 DR EMBL; AB017903; BAA33469.1; -;
 DR EMBL; AB017904; BAA33470.1; -;
 DR InterPro; IPR003461; Keratin.
 DR Pfam; PF02422; Keratin; 1.
 KW Keratin; Multigene family; Feather; Acetylation.
 FT INIT MET 0
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 100 AA; 10235 MW; 54292E196A53DF99 CRC64;
 Alignment Scores:
 Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x KRF1_COLLI (1-100)
 QY 2512 CTGAGCTGCATTACTCTAGA 2532
 Db 85 LeuSerCysIleThrSerArg 91
 RESULT 60
 KRF2_COLLI
 ID KRF2_COLLI STANDARD; PRT; 100 AA.
 AC Q93439;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Feather keratin Cosl-2 (F-ker).
 OS Columba livia (Domestic pigeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
 OX NCBI_TaxID=8932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Takahashi R.;
 RT "Pigeon feather keratin genes.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A
 CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB017902; BAA33468.1; -;
 DR InterPro; IPR003461; Keratin.
 DR Pfam; PF02422; Keratin; 1.
 KW Keratin; Multigene family; Feather; Acetylation.
 FT INIT MET 0
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 100 AA; 10166 MW; 53C50756254E5CED CRC64;
 Alignment Scores:
 Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0

Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF2_COLLI (1-100)

QY 2512 CTGAGCTGCATTACCTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 61

KRF3_COLLI
 ID KRF3_COLLI STANDARD; PRT; 100 AA.
 AC O93500;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Feather keratin Cos2-2 (F-ker).
 OS Columba livia (Domestic pigeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columba.
 OX NCBI_TaxID=8932;
 RP SEQUENCE FROM N.A.
 RA Takahashi R.;

DR EMBL; AB017905; BAA33471.1; -;
 DR InterPro; IPR003461; Keratin.
 DR Pfam; PF02422; Keratin; 1.
 KW Keratin; Multigene family; Feather; Acetylation.
 FT INIT MET 0 BY SIMILARITY
 FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 100 AA; 10197 MW; A58E713B92D9E065 CRC64;
 Alignment Scores:
 Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF3_COLLI (1-100)
 QY 2512 CTGAGCTGCATTACCTCTAGA 2532

Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF3_COLLI (1-100)

QY 2512 CTGAGCTGCATTACCTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 62

KRF4_COLLI
 ID KRF4_COLLI STANDARD; PRT; 100 AA.
 AC Q9PSV3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Feather keratin Cos2-3 (F-ker).
 OS Columba livia (Domestic pigeon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Columbiformes; Columba.
 OX NCBI_TaxID=8932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takahashi R.;
 RT "Pigeon feather keratin genes.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A
 CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB017906; BAA33472.1; -;
 DR InterPro; IPR003461; Keratin.
 DR Pfam; PF02422; Keratin; 1.
 KW Keratin; Multigene family; Feather; Acetylation.
 FT INIT MET 0 BY SIMILARITY
 FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
 SQ SEQUENCE 100 AA; 10288 MW; BB885DF31EE0952F CRC64;

Alignment Scores:
 Pred. No.: 268 Length: 100
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF4_COLLI (1-100)

QY 2512 CTGAGCTGCATTACCTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 63

YC54_AQUAE
 ID YC54_AQUAE STANDARD; PRT; 102 AA.
 AC O67255;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AQ_1254.
 GN AQ_1254.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., AuJay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).

US-09-774-490-1 (1-2709) x KRF4_COLLI (1-100)

CC -!- SIMILARITY: BELONGS TO THE UPF0092 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE000731; AAC07266.1; -
DR PIR; D70408; D70408; YajC.
DR InterPro; IPR003849; YajC.
DR Pfam; PF02699; YajC; 1.
DR TIGRFAMs; TIGR00739; YajC; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
SQ SEQUENCE 102 AA; 11823 MW; 2361BFC80B2F3277 CRC64;
Alignment Scores:
Pred. No.: 268 Length: 102
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x YC54_AQUAE (1-102)
QY 1404 RAGGTATATACATCATCAGGA 1384
DB 51 LysValIleThrSerSerGly 57
RESULT 64
ID INS2_XENLA STANDARD; PRT; 106 AA.
AC P12707;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin 2 precursor.
GN INS2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89255444; PubMed=2722842;
RA Shuldiner A.R., Phillips S., Roberts C.T. Jr., Lerioth D., Roth J.;
RT "Xenopus laevis contains two nonallelic preproinsulin genes. cDNA
cloning and evolutionary perspective.";
RL J. Biol. Chem. 264:9428-9432(1989).
RN [2]
RP SEQUENCE OF 24-53 AND 86-106.
RX MEDLINE=89289601; PubMed=2661211;
RA Shuldiner A.R., Bennett C., Robinson E.A., Roth J.;
RT "Isolation and characterization of two different insulins from an
amphibian, Xenopus laevis";
RL Endocrinology 125:469-477(1989).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24442; AAA49887.1; -
DR PIR; B33847; IPXL2.
DR HSSP; P01315; 1SDB.
DR InterPro; IPR004825; Ins/IGF/relax.

DR Pfam; PF00049; Insulin; 1.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Multigene family; Signal.
FT SIGNAL 1 23
FT CHAIN 24 53 INSULIN 2 B CHAIN.
FT PROPEP 56 83 INSULIN 2 C PEPTIDE.
FT CHAIN 86 106 INSULIN 2 A CHAIN.
FT DISULFID 30 92 INTERCHAIN.
FT DISULFID 42 105 INTERCHAIN.
FT DISULFID 91 96
SQ SEQUENCE 106 AA; 12207 MW; 15859A0C758E7DEF CRC64;
Alignment Scores:
Pred. No.: 267 Length: 106
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x INS2_XENLA (1-106)
QY 1977 GTGCTACTATTCTCTACACCA 1957
DB 13 ValLeuLeuPheSerThrPro 19
RESULT 65
ID Y094_HAEIN STANDARD; PRT; 106 AA.
AC P43939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0094.
GN HI0094.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32695; AAC21779.1; -
DR PIR; B64001; B64001.
DR TIGR; HI0094; -
DR InterPro; IPR003474; Glcn transporter.
DR Pfam; PF02447; GntP_permease; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 78 98 POTENTIAL.
SQ SEQUENCE 106 AA; 11735 MW; 672EC13E1F59E3E4 CRC64;

Alignment Scores:
 Pred. No.: 267
 Score: 7.00
 Length: 106
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Indels: 0
 Query Match: 0.80%
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x Y094_HAEIN (1-106)

Qy 283 ATTGTTCTCCCATCTGATA 263

Db 89 llevalleuProileLeuLys 95

RESULT 66

LHG2_RHOCA

ID_LHG2_RHOCA STANDARD; PRT; 118 AA.

AC P23460;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 01-NOV-1991 (Rel. 20, Last annotation update)

DE Light-harvesting protein B-800/850, gamma chain.

GN PUCR.

OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;

OC Rhodobacteraceae; Rhodobacter.

OX NCBI_TaxID=1061;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89359127; PubMed=2549005;

RA Tichy H.V., Oberle B., Stiehle H., Schiltz E., Drews G.;

RT "Genes downstream from pucB and pucA are essential for formation of

the B800-850 complex of Rhodobacter capsulatus.";

RL J. Bacteriol. 171:4914-4922(1989)

CC -!- FUNCTION: SEEMS TO BE REQUIRED FOR THE LH-II STABILIZATION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M28510; AAA26165.1; -.

DR PIR; E33958; E33958.

KW Antenna complex; Light-harvesting polypeptide.

SQ SEQUENCE 118 AA; 13493 MW; 749AF4CC1EF82262 CRC64;

Alignment Scores:

Pred. No.: 264
 Score: 7.00
 Length: 118
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Query Match: 0.80%
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x LHG2_RHOCA (1-118)

Qy 1685 TGGAAAGCTCCATCTGCTGAAA 1665

Db 84 TrpLysAlaProLeuLeuLys 90

RESULT 67

CRCB_NEIMA

ID_CRCB_NEIMA STANDARD; PRT; 119 AA.

AC Q9JUL1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Protein crCB homolog.

GN CRCB OR NNA1264.

OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=65699;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z2491 / Serogroup A / Serotype 4A;

RX MEDLINE=20222556; PubMed=10761919;

RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,

RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,

RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,

RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,

RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,

RA Whitehead S., Spratt B.G., Barrrell B.G.;

RT "Complete DNA sequence of a serogroup A strain of Neisseria

meningitidis Z2491.";

RL Nature 404:502-506(2000).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE CRCB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AL162755; CAB84519.1; -.

DR PIR; F81894; F81894.

DR HAWAP; MF_00454; -; 1.

DR InterPro; IPR003691; Camphor_Crcb.

PFam; PF02537; CRCB; 1.

KW Transmembrane; Complete proteome.

FT TRANSMEM 5 22 POTENTIAL.

FT TRANSMEM 37 54 POTENTIAL.

FT TRANSMEM 61 83 POTENTIAL.

FT TRANSMEM 93 112 POTENTIAL.

SQ SEQUENCE 119 AA; 12456 MW; D3104F0B06F5BD77 CRC64;

Alignment Scores:

Pred. No.: 264
 Score: 7.00
 Length: 119
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Query Match: 0.80%
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x CRCB_NEIMA (1-119)

Qy 1167 ACTGGATTTTGGAGTCTTTA 1147

Db 65 ThrGlyPheLeuGlySerLeu 71

RESULT 68

CRCB_NEIMB

ID_CRCB_NEIMB STANDARD; PRT; 119 AA.

AC Q9JZG6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Protein crCB homolog.

GN CRCB OR NMB1065.

OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MC58 / Serogroup B;

RX MEDLINE=20175755; PubMed=10710307;

RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cirtone H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vanathavan J.,
 RA Gill J., Scariato V., Maignani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE CRCB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE002457; AAF41461.1; -;
 DR PIR; C81125; C81125.
 DR TIGR; NMB1065; -;
 DR HAMAP; MF_00454; -; 1.
 DR InterPro; IPR003691; Camphor_CrcB.
 DR Pfam; PF02537; CRCB; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 5 22 POTENTIAL.
 FT TRANSMEM 37 54 POTENTIAL.
 FT TRANSMEM 61 83 POTENTIAL.
 FT TRANSMEM 93 112 POTENTIAL.
 SQ SEQUENCE 119 AA; 12474 MW; D651C88215B9B80B CRC64;
 Alignment Scores:
 Pred. No.: 264 Length: 119
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x CRCB_NEIMB (1-119)
 QY 1167 ACTGGATTTTAGGATCTTTA 1147
 Db 65 ThrGlyPheLeuGlySerLeu 71
 RESULT 69
 CRCB_PASMU
 ID CRCB_PASMU STANDARD; PRT; 128 AA.
 AC Q9CKT7.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Protein crCB homolog.
 GN CRCB OR PM1630.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE CRCB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE006200; AAK03714.1; ALT_INIT.
 DR HAMAP; MF_00454; -; 1.
 DR InterPro; IPR003691; Camphor_CrcB.
 DR Pfam; PF02537; CRCB; 1.
 DR TIGR; TIGR00494; CrcB; 1.
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 7 26 POTENTIAL.
 FT TRANSMEM 36 58 POTENTIAL.
 FT TRANSMEM 70 92 POTENTIAL.
 FT TRANSMEM 102 124 POTENTIAL.
 SQ SEQUENCE 128 AA; 14139 MW; 13D3280EA5A1E805 CRC64;
 Alignment Scores:
 Pred. No.: 262 Length: 128
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x CRCB_PASMU (1-128)
 QY 1167 ACTGGATTTTAGGATCTTTA 1147
 Db 74 ThrGlyPheLeuGlySerLeu 80
 RESULT 70
 ATP6_THUOB
 ID ATP6_THUOB STANDARD; PRT; 133 AA.
 AC Q36090;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ATP synthase A chain (EC 3.6.3.14) (Protein 6) (Fragment).
 GN MTATP6 OR ATP6.
 OS *Thunnus obesus* (Bigeye tuna).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
 OC Scombridae; Thunnus.
 OX NCBI_TaxID=8241;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96139007; PubMed=8587119;
 RA Chow S., Kishino H.;
 RT "Phylogenetic relationships between tuna species of the genus *Thunnus*
 RT (*Scombridae*; *Teleostei*): inconsistent implications from morphology,
 RT nuclear and mitochondrial genomes.";
 RL J. Mol. Evol. 41:741-748(1995).
 CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
 CC DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC


```
-----
CC EMBL; D63416; BAA09721.1;
DR InterPro; IPR000568; ATP-Synt Asub.
DR Pfam; PF00119; ATP-Synt A; 1.
DR PRINTS; PR00123; ATPASEA.
DR PROSITE; PS00449; ATPASE_A; PARTIAL.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT NON TER 1
FT NON TER 133
SQ SEQUENCE 133 AA; 14982 MW; D625359D2374084B CRC64;

Alignment Scores:
Pred. No.: 262 Length: 133
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x ATP6_THUOB (1-133)
OY 1248 CTTCTCATCACTCATGCTA 1228
Db 51 LeuLeuThrSerLeuMetIeu 57

RESULT 71
SRP_CHLPS
ID SRP_CHLPS STANDARD; PRT; 134 AA.
AC P28164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sulfur-rich protein.
GN SRP.
OS Chlamydia psittaci (Chlamydophila psittaci).
OC Bacteriia; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83554;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=6BC;
RX MEDLINE=91267949; PubMed=2050637;
RA Everett K.D.E., Hatch T.P.;
RT "Sequence analysis and lipid modification of the cysteine-rich envelope proteins of Chlamydia psittaci 6BC.";
RL J. Bacteriol. 173:3821-3830(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61116; AAB61620.1;
SQ SEQUENCE 134 AA; 14732 MW; 2982C6FC51BF854 CRC64;

Alignment Scores:
Pred. No.: 261 Length: 134
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SRP_CHLPS (1-134)
OY 592 GCGGTACAGTGCAGTCTGATT 572
Db 29 GlyValGlnValSerLeuIle 35

RESULT 72
YBJ5_CANAL
```

```
YBJ5_CANAL STANDARD; PRT; 134 AA.
ID PS3716;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in WHS11 5'region (Fragment).
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=WO-1;
RX MEDLINE=93380673; PubMed=7916716;
RA Srikantha T., Soll D.R.;
RT "A white-specific gene in the white-opaque switching system of Candida albicans.";
RL Gene 131:53-60(1993).
RN 2
RP CONCEPTUAL TRANSLATION.
RA Bairoch A.;
RL Unpublished observations (AUG-1996).
CC 1- SIMILARITY: STRONG, TO YEAST YBL0835.
CC 1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; FRAMESHIFTS WERE INTRODUCED TO MAXIMIZE THE SIMILARITY WITH YEAST YBL095W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65451; -; NOT_ANNOTATED_CDS.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 134 AA; 14778 MW; 58AE66F982E866FA CRC64;

Alignment Scores:
Pred. No.: 261 Length: 134
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YBJ5_CANAL (1-134)
OY 728 ACAGCATCCCTTTTAAATAGAT 748
Db 62 ThrAlaSerLeuLeuIleAsp 68

RESULT 73
UGR1_MOUSE
ID UGR1_MOUSE STANDARD; PRT; 139 AA.
AC Q920H1; Q920H2; Q920H3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uteroglobin-related protein 1 precursor (Secretoglobulin family 3A member 2).
GN SCGB3A2 OR UGRP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC TISSUE=Lung;
RX MEDLINE=21539178; PubMed=11682631;
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C., Kimura S.;
```

RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is
RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1
RL Mol. Endocrinol. 15:2021-2036(2001).
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=alternative splicing; Named isoforms=3;
CC Name=C;
CC IsoId=Q920H1-1; Sequence=Displayed;
CC Name=A;
CC IsoId=Q920H1-2; Sequence=VSP_006727, VSP_006728;
CC Name=B;
CC IsoId=Q920H1-3; Sequence=VSP_006726;
CC -!- TISSUE SPECIFICITY: Highly expressed in lung.
CC -!- SIMILARITY: BELONGS TO THE UTEROGLIBIN FAMILY. UGRP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF274959; RAL25708.1; -;
CC EMBL; AF274960; RAL25709.1; -;
CC EMBL; AF274961; RAL25710.1; -;
CC MGD; MG1:2153470; Scp33a2.
CC GO; GO:0005576; C:extracellular; IDA.
CC GO; GO:0005515; F:protein binding activity; IPI.
CC InterPro; IPR006038; Uteroglobin_supf.
CC Pfam; PF01099; Uteroglobin; 1.
KW Signal; Alternative splicing. POTENTIAL.
FT SIGNAL 1 21
FT CHAIN 22 139 UTEROGLIBIN-RELATED PROTEIN 1.
FT VARSPLIC 107 139 VSVLFLPMICAYPRDSKKQTAFIERVFEQSKL-> EALS
FT
FT
FT VARSPLIC 85 91 /FTID=VSP_006726.
FT
FT VARSPLIC 92 139 /FTID=VSP_006727.
FT
FT /FTID=VSP_006728.
SQ SEQUENCE 139 AA; 15431 MW; 8A2FB080B41E65E4 CRC64;

Alignment Scores:
Pred. No.: 261 Length: 139
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x UGR1_MOUSE (1-139)

QY 1212 GATCCCTTGAAATGCTACTG 1192
|||||
DB 45 AspProLeuLysMetLeuLeu 51

RESULT 74
HBA5_XENLA
ID_HBA5_XENLA STANDARD; PRT; 141 AA.
AC P06638;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-5 chain (Alpha-TS).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxid=8355;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=85297748; PubMed=2993998;
RA Barville D.; Williams J.G.;
RT "The pattern of expression of the Xenopus laevis tadpole alpha-globin
RT genes and the amino acid sequence of the three major tadpole alpha-
RT globin polypeptides.";
RL Nucleic Acids Res. 13:5407-5421(1985).
CC -!- FUNCTION: THIS IS A LARVAL (TADPOLE) ALPHA-GLOBIN.
CC -!- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- SIMILARITY: Belongs to the globin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02798; CAA26566.1; -;
CC PIR; C24338; C24338.
CC HSRP; P01922; IABW.
CC InterPro; IPR002338; Alpha haem.
CC Pfam; PF00042; globin; 1.
CC PRINTS; PR00612; ALPHAHAE.
CC PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT INIT_MET 0 0
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15270 MW; 73E02DE6BF6CEB CRC64;

Alignment Scores:
Pred. No.: 260 Length: 141
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x HBA5_XENLA (1-141)

QY 1346 TGGCACAAGTTCCTGGCCGTG 1326
|||||
DB 125 TrpAspLysPheLeuAlaVal 131

RESULT 75
PSG8_HUMAN
ID_PSG8_HUMAN STANDARD; PRT; 144 AA.
AC Q9UQ74;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pregnancy-specific beta-1-glycoprotein 8 precursor (PSBG-8)
DE (Fragment).
GN PSG8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Frangemyr L.; Teglund S.; Israelsson A.; Hammarstrom S.;
RT "Characterization of upstream promoter region, exon1 and exon2 of the
RT PSB gene family.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- DEVELOPMENTAL STAGE: PSBG ARE PRODUCED IN HIGH QUANTITY DURING
CC PREGNANCY.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
CC SUBFAMILY.

```
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF106556; AAD21022.1; -
DR EMBL; AF106555; AAD21022.1; JOINED.
DR Genew; HGNC:9525; PSG8.
DR MIM; 176397; -
DR InterPro; IPR003599; IG.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; FALSE NEG.
KW Immunoglobulin domain; Glycoprotein; Signal; Repeat; Multigene family.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 >144 PREGNANCY-SPECIFIC BETA-1-GLYCOPROTEIN 8.
FT DOMAIN 35 144 IG-LIKE V-TYPE.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 16109 MW; 0A83A6711042C1FE CRC64;

Alignment Scores:
Pred. No.: 260 Length: 144
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PSG8_HUMAN (1-144)
Qy 722 CTGCTGACAGCATCCCTTTTA 742
Db 19 LeuLeuThrAlaSerLeuLeu 25
```

Search completed: August 3, 2003, 10:32:23
Job time : 127 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:37 ; Search time 128.5 Seconds
(without alignments)
10880.381 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatcttttttttcgatg.....aggcttttttttcctaataacc 2709

Scoring table: OLIGO Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 830525 seqs, 258052604 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1660916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlh
-Q=/csn2.1/USPTO/spool/US09774490/runat_03082003_102927_4459/app_query.fasta_1.2887
-DB=SPTRMBL_23 -QWMT=fastan -SUFFIX=oligo.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CCN 1 1 0 @runat_03082003_102927_4459 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	43	4.8	774	13 Q8JIW9	Q8JIW9 xenopus lae

Q9ns98	homo sapien	782	4	Q9NS98	Q9ns98
Q9hbr1	homo sapien	15	4	Q9HBR1	Q9hbr1
Q96gx0	homo sapien	15	4	Q96GX0	Q96gx0
Q8tb71	homo sapien	15	4	Q8TB71	Q8tb71
Q8tdv7	homo sapien	15	4	Q8TDV7	Q8tdv7
Q8bkq6	mus musculus	14	11	Q8BKQ6	Q8bkq6
Q8bm6	mus musculus	777	11	Q8BMF6	Q8bm6
Q8bh34	mus musculus	777	11	Q8BH34	Q8bh34
Q9yhx3	brachydanio	10	13	Q9YHX3	Q9yhx3
Q8ncq1	homo sapien	13	15	Q8NCQ1	Q8ncq1
Q8qgu9	gallus gall	12	15	Q8QGU9	Q8qgu9
Q9qz23	mus musculus	13	15	Q9QZ23	Q9qz23
Q8nas4	homo sapien	14	12	Q8NAS4	Q8nas4
Q9ji29	rattus norv	15	12	Q9JI29	Q9ji29
Q8axa8	gallus gall	16	10	Q8AXA8	Q8axa8
Q64906	alcelaphine	17	10	Q64906	Q64906
Q9c0c4	homo sapien	10	11	Q9C0C4	Q9c0c4
Q96qn2	homo sapien	139	4	Q96QN2	Q96qn2
Q49704	mycobacteri	223	2	Q49704	Q49704
Q54948	mus musculus	284	11	Q54948	Q54948
Q33142	mycobacteri	311	16	Q33142	Q33142
Q9bw98	homo sapien	325	4	Q9BW98	Q9bw98
Q8nbn9	homo sapien	333	4	Q8NB9	Q8nbn9
Q8f7r6	leptospira	350	16	Q8F7R6	Q8f7r6
Q8wvp9	homo sapien	458	4	Q8WVP9	Q8wvp9
Q96fk5	homo sapien	475	4	Q96FK5	Q96fk5
Q9nx92	homo sapien	510	4	Q9NX92	Q9nx92
Q8bi3	mus musculus	550	11	Q8BIC3	Q8bi3
Q8bjc1	mus musculus	799	11	Q8BJC1	Q8bjc1
Q8bif6	mus musculus	823	11	Q8BIF6	Q8bif6
Q9c0b8	homo sapien	893	4	Q9C0B8	Q9c0b8
Q8r3m7	mus musculus	911	11	Q8R3M7	Q8r3m7
Q8r1b4	mus musculus	911	11	Q8R1B4	Q8r1b4
Q95201	mustela vis	4859	6	Q95201	Q95201
Q91319	rana catesb	4868	13	Q91319	Q91319
Q90985	gallus gall	4869	13	Q90985	Q90985
Q9t633	oryctolagus	4872	6	Q9TS33	Q9t633
Q9ern6	mus musculus	4967	11	Q9ERN6	Q9ern6
Q29621	oryctolagus	4968	6	Q29621	Q29621
Q91313	rana catesb	5037	13	Q91313	Q91313
Q13054	makaira nig	5081	13	Q13054	Q13054
Q24498	drosophila	5107	5	Q24498	Q24498
Q8mks5	drosophila	5113	5	Q8MK53	Q8mks5
Q8mks4	drosophila	5127	5	Q8MK54	Q8mks4
Q8k4m4	mus musculus	48	11	Q8K4M4	Q8k4m4
Q924C2	shigella so	51	2	Q924C2	Q924C2
Q9335	pyrococcus	101	17	Q9335	Q9335
Q8ze0	versinia pe	178	16	Q8ZE0	Q8ze0
Q9s787	arabidopsis	180	10	Q9S787	Q9s787
Q9mb00	brassica ca	184	10	Q9MB00	Q9mb00
Q8cl36	versinia pe	218	16	Q8CL36	Q8cl36
Q9kam9	bacillus ha	223	16	Q9KAM9	Q9kam9
Q8hm61	lophius ame	227	8	Q8HM61	Q8hm61
Q96id9	homo sapien	277	4	Q96ID9	Q96id9
Q9v3f8	drosophila	280	5	Q9V3F8	Q9v3f8
Q96643	drosophila	280	5	Q96643	Q96643
Q22841	arabidopsis	283	10	Q22841	Q22841
Q91632	xenopus lae	292	13	Q91632	Q91632
Q91633	xenopus lae	292	13	Q91633	Q91633
Q8d3m7	vibrio vuln	294	16	Q8D3M7	Q8d3m7
Q8d3q3	vibrio vuln	299	16	Q8D3Q3	Q8d3q3
Q9a6s1	caulobacter	309	16	Q9A6S1	Q9a6s1
Q8nh68	homo sapien	312	4	Q8NH68	Q8nh68
Q97d40	clostridium	326	16	Q97D40	Q97d40
Q91914	xenopus lae	327	13	Q91914	Q91914
Q97d45	clostridium	350	16	Q97D45	Q97d45
Q9bye7	homo sapien	352	4	Q9BYE7	Q9bye7
Q99na9	mus musculus	353	11	Q99NA9	Q99na9
Q90753	feline herp	371	12	Q90753	Q90753
Q9aqp4	pseudomonas	386	2	Q9AQP4	Q9aqp4
Q8gi39	pseudomonas	386	2	Q8GI39	Q8gi39
Q8rth6	vibrio chol	402	2	Q8RTH6	Q8rth6

C 75 8 0.9 409 2 052274
 C 76 8 0.9 421 5 09XUV7
 C 77 8 0.9 424 2 093RK2
 C 78 8 0.9 428 2 063759
 C 79 8 0.9 437 4 08N7U5
 C 80 8 0.9 449 16 08EFM5
 C 81 8 0.9 475 16 092YE1
 C 82 8 0.9 485 5 09V9N1
 C 83 8 0.9 485 16 09CLJ7
 C 84 8 0.9 519 2 08RLN3
 C 85 8 0.9 521 2 093MK3
 C 86 8 0.9 524 2 093T64
 C 87 8 0.9 524 2 0938U2
 C 88 8 0.9 534 16 09EWX5
 C 89 8 0.9 544 16 08PMW3
 C 90 8 0.9 571 16 092BB4
 C 91 8 0.9 571 16 09Y6T8
 C 92 8 0.9 585 10 042880
 C 93 8 0.9 648 8 09TKV7
 C 94 8 0.9 664 16 08XL36
 C 95 8 0.9 671 6 09SLG6
 C 96 8 0.9 737 5 09TYX1
 C 97 8 0.9 758 17 08PW83
 C 98 8 0.9 761 4 08WUA9
 C 99 8 0.9 770 13 093394
 C 100 8 0.9 810 16 08ET53

ALIGNMENTS

RESULT 1
 Q8JIW9 PRELIMINARY; PRT; 774 AA.
 ID Q8JIW9
 AC Q8JIW9
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Semaphorin 3A.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Head;
 RA Tannahill D., Nielsen J., Regan A.G.;
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AV030051; AAK38166.1; -;
 DR EMBL; AV030051; AAK38166.1; -;
 DR InterPro; IPR003599; Ig-like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR001005; MYB DNA binding.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00037; MYB; 1.
 SQ SEQUENCE 774 AA; 89319 MW; 3127FD2E0CCD5FA4 CRC64;

Alignment Scores:
 Pred. No.: 1.83e-37 Length: 774
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.80% Indels: 0
 DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q8JIW9 (1-774)
 QY 1004 TGCAGATGACTTGGAGGGCACAGAGTCTGGTAATATGACACATTCCTCAA 1063
 DB 270 CyslysAsnAspPheGlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLys 289
 QY 1064 GCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCATTGACACTCATTTTGGATGAAC 1123
 DB 290 AlaArgLeuLeuCysSerValProGlyProAsnGlyIleAspThrHisPheAspGluLeu 309
 QY 1124 CAGGATGTA 1132
 DB 310 GlnAspVal 312
 RESULT 2
 Q9NS98 PRELIMINARY; PRT; 782 AA.
 ID Q9NS98
 AC Q9NS98; Q9H7Q3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Semaphorin sem2 (FLJ00014 protein).
 GN SEM2 OR FLJ00014.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayaashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.;
 RT "Human semaphorin";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 58-782 FROM N.A.
 RC TISSUE-Spleen;
 RA Chara O., Nagase T., Kikuno R., Okumura K.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human spleen";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB029496; BAA98132.1; -;
 DR EMBL; AK024425; BAB15715.1; -;
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;
 Alignment Scores:
 Pred. No.: 8.64e-09 Length: 782
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: 4 Gaps: 0
 US-09-774-490-1 (1-2709) x Q9NS98 (1-782)
 QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGTGGGTGGTGT 1816
 DB 528 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544
 RESULT 3
 Q9HBR1 PRELIMINARY; PRT; 457 AA.
 ID Q9HBR1
 AC Q9HBR1

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083186; BAB88870.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 754 AA; 83691 MW; E61FD08C04E9A68E CRC64;
Alignment Scores:
Pred. No.: 1.4e-06 Length: 754
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
Gaps: 0
DB: 4
US-09-774-490-1 (1-2709) x Q8TDV7 (1-754)
Qy 1028 AGAAGCTCTGGTGAATAAATGGACAACTTCCTCAAGCTCGTCTG 1072
Db 277 ARGSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 291
RESULT 7
Q8BKQ6
ID Q8BKQ6 PRELIMINARY; PRT; 403 AA.
AC Q8BKQ6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK051165; BAC34542.1; -
SQ SEQUENCE 403 AA; 45575 MW; 1D5B284549825884 CRC64;
Alignment Scores:
Pred. No.: 1.93e-05 Length: 403
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
DB: 11
US-09-774-490-1 (1-2709) x Q8BKQ6 (1-403)
Qy 1040 AATAAATGGACAACTTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 8
Q8BMF6
ID Q8BMF6 PRELIMINARY; PRT; 777 AA.
AC Q8BMF6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK031704; BAC27522.1; -
SQ SEQUENCE 777 AA; 89562 MW; B7204D82288B89CE CRC64;
Alignment Scores:
Pred. No.: 1.76e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
DB: 11
US-09-774-490-1 (1-2709) x Q8BMF6 (1-777)
Qy 1040 AATAAATGGACAACTTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 9
Q8BH34
ID Q8BH34 PRELIMINARY; PRT; 777 AA.
AC Q8BH34;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney, and Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK028900; BAC26185.1; -
DR EMBL; AK052671; BAC35092.1; -
SQ SEQUENCE 777 AA; 89548 MW; 5450D8D45D1BDABF CRC64;
Alignment Scores:
Pred. No.: 1.76e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
DB: 11
US-09-774-490-1 (1-2709) x Q8BH34 (1-777)
Qy 1040 AATAAATGGACAACTTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 10
Q9VHX3

ID Q9YHX3 PRELIMINARY; PRT; 328 AA.
 AC Q9YHX3;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Secreted class III semaphorin (Fragment).
 GN SEMA8 OR SEMA28
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99060909; PubMed=9867349;
 RA Halloran M.C., Severance S.M., Yee C.S., Genza D.L., Kuwada J.Y.;
 RT "Molecular cloning and expression of two novel zebrafish
 RT semaphorins.";
 RL Mech. Dev. 76:165-168(1998).
 DR EMBL; AF073290; AAC72346.1; -.
 DR ZFIN; ZDB-GENE-990715-8; sema8.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 DR NON TER 1
 FT SEQUENCE 328 AA; 37375 MW; 9AB12F43DE28BA82 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 0.000252 Length: 328
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 13 Gaps: 0
 US-09-774-490-1 (1-2709) x Q9YHX3 (1-328)
 QY 1766 CGTGTGCTGAGTGTGCTGCCGAGACCCCTTACTGT 1804
 Db 94 AlacysAlaGluCysCysLeuAlaArgAspProTyrCys 106
 RESULT 11
 ID Q8NCQ1 PRELIMINARY; PRT; 416 AA.
 AC Q8NCQ1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC029590; AAH29590.1; -.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; IG; 1.
 DR SMART; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 756 AA; 85196 MW; C36754C02541ED88 CRC64;
 Alignment Scores:
 Pred. No.: 0.000224 Length: 756
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 13 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8QGU9 (1-756)
 QY 1778 TGTTCCTCGCCGAGACCCCTTACTGTGCTGGATGGT 1816
 Db 527 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 539
 RESULT 13
 ID Q9QX23 PRELIMINARY; PRT; 775 AA.
 AC Q9QX23;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DR PROSITE; PS50835; IG LIKE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 416 AA; 48691 MW; 61D92D89CAE70631 CRC64;
 Alignment Scores:
 Pred. No.: 0.000244 Length: 416
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 4 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8NCQ1 (1-416)
 QY 1778 TGTTCCTCGCCGAGACCCCTTACTGTGCTGGATGGT 1816
 Db 183 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 195
 RESULT 12
 ID Q8QGU9 PRELIMINARY; PRT; 756 AA.
 AC Q8QGU9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin 3F.
 GN SEMA3F.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matanabe Y., Nakamura H.;
 RT "Axon guidance of the trochlear nerve by Sema 3F along mid-hindbrain
 RT boundary.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB072930; BAB88691.1; -.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG C2.
 DR InterPro; IPR003006; IG MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; IG; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00408; IG C2; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Immunoglobulin domain.
 SQ SEQUENCE 756 AA; 85196 MW; C36754C02541ED88 CRC64;
 Alignment Scores:
 Pred. No.: 0.000224 Length: 756
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 13 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8QGU9 (1-756)
 QY 1778 TGTTCCTCGCCGAGACCCCTTACTGTGCTGGATGGT 1816
 Db 527 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 539
 RESULT 13
 ID Q9QX23 PRELIMINARY; PRT; 775 AA.
 AC Q9QX23;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin M-Semak.
 GN SEMA3E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/Black 6;
 RA Miyazaki N., Furiyama T., Inagaki S.;
 RT "A novel semaphorin, M-Semak which inhibits neural outgrowth from
 RT sensory neurons."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF034744; RAD01996.1; -
 DR MGD; MGI:1340034; SEMA3E.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG.MHC.
 DR InterPro; IPR003659; plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 SQ SEQUENCE 775 AA; 89543 MW; 221E766FF404098D4 CRC64;

Alignment Scores:
 Pred. No.: 0.000223 Length: 775
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q90X23 (1-775)

QY 1778 TGTTCCTCGCCGAGACCTTACTGCTGGTGGATGGT 1816
 Db 529 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 14
 Q8NAS4 PRELIMINARY; PRT; 132 AA.
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ34863.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furiyama T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamiyama K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matauo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK092182; BAC03823.1; -
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 132 AA; 14227 MW; BF1D942742DE0FA1 CRC64;

Alignment Scores:
 Pred. No.: 0.00363 Length: 132
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.34% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8NAS4 (1-132)

QY 1547 TTTATCGAACAGATGTTGGACCGTCTTAAAGTA 1582
 Db 42 PheileGlyThrAspValGlyThrValLeuLysVal 53

RESULT 15
 Q9J129 PRELIMINARY; PRT; 296 AA.
 ID Q9J129
 AC Q9J129;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Brenz Verca M.S., Widmer D.A.J., Wagner G.C., Dreyer J.L.;
 RT "New rat semaphorin with closest homology to avian Sema3D (C-
 RT collapsein-2)."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF268594; AAF76329.1; -
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR NON_TER 1
 FT NON_TER 296
 SQ SEQUENCE 296 AA; 33558 MW; C2226212885D044C CRC64;

Alignment Scores:
 Pred. No.: 0.00324 Length: 296
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.34% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q9J129 (1-296)

QY 1319 CCCTATCCAGCCGAGCACTGTCCGAGCAACACA 1354
 Db 136 ProTyrProArgProGlyThrCysProSerLysThr 147

RESULT 16
 Q8AXA8 PRELIMINARY; PRT; 160 AA.
 ID Q8AXA8
 AC Q8AXA8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Secreted chemorepellent semaphorin 3F (Fragment).
 GN SEMA3F.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chilton J.K., Guthrie S.;
 RT "Cranial expression of semaphorins and neuropilins in the embryonic

RT chick.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF495525; AAN77007.1; -;
 FT NON TER 160
 SQ SEQUENCE 160 AA; 17865 MW; 8033FOC191EAF2DE CRC64;

Alignment Scores:
 Pred. No.: 0.567 Length: 160
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q9AXA8 (1-160)

QY 1106 ACTATTTTGATGACTGAGGATGATTC 1135
 |||||
 Db 108 ThrHisPheAspGluLeuGlnaspValPhe 117

RESULT 17
 Q64906 PRELIMINARY; PRT; 653 AA.
 AC Q64906;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to GENBANK ACCESSION number L26081.
 OS Alcelaphine herpesvirus 1 (wildbeest herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=35252;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RX MEDLINE=97201573; PubMed=9049359;
 RA Enser A.; Fleckenstein B.;
 RT "Alcelaphine herpesvirus type 1 has a semaphorin-like gene."
 RL J. Gen. Virol. 76:1063-1067(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RX MEDLINE=97404659; PubMed=9261371;
 RA Enser A.; Pflanz R.; Fleckenstein B.;
 RT "Primary structure of the alcelaphine herpesvirus 1 genome."
 RL J. Virol. 71:6517-6525(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RA Enser A.; Pflanz R.; Fleckenstein B.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U18243; AAC54475.1; -;
 DR EMBL; AF005370; AAC58054.1; -;
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 SQ SEQUENCE 653 AA; 73645 MW; 5C2816BD567E706E CRC64;

Alignment Scores:
 Pred. No.: 0.466 Length: 653
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 12 Gaps: 0

US-09-774-490-1 (1-2709) x Q64906 (1-653)

QY 1043 AATCGACACATTTCTCAAGCTGCTG 1072
 |||||
 Db 302 LysTrpThrPheLeuLysAlaArgLeu 311

RESULT 18
 Q9C0C4 PRELIMINARY; PRT; 963 AA.
 ID Q9C0C4;
 AC Q9C0C4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein KIAA1739 (Fragment).
 GN KIAA1739.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T.; Kikuno R.; Hattori A.; Kondo Y.; Okumura K.; Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 7:347-355(2000).
 DR EMBL; AB051526; BAB21830.1; -;
 DR Genew; HGNC:10731; SEMA4C.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 963 AA; 106735 MW; 3FA37DBBA3483ECE CRC64;

Alignment Scores:
 Pred. No.: 0.441 Length: 963
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9C0C4 (1-963)

QY 917 GATGACAAAGTATCTTTTCTCCGTGAA 946
 |||||
 Db 362 AspaAspLysValTyrPhePheArgGlu 371

RESULT 19
 Q96QN2 PRELIMINARY; PRT; 139 AA.
 ID Q96QN2;
 AC Q96QN2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Complement component 1, s subcomponent.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 DR EMBL; BC007010; AAH07010.1; -;
 DR InterPro; IPR000859; CUB_domain.

DR Pfam; PF00431; CUB; 1.
 DR SMART; SM00042; CUB; 1.
 DR PROSITE; PS01180; CUB; 1.
 SQ SEQUENCE 139 AA; 15969 MW; 590D59BC718112D5 CRC64;

Alignment Scores:
 Pred. No.: 7.33 Length: 139
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96QN2 (1-139)

QY 151 TTGTGGCGGCAGAGAGTTCAAACAT 125
 Db 82 LeuCySGlyGlnArgSerSerAsnAsn 90

RESULT 20

Q49704 PRELIMINARY; PRT; 223 AA.
 AC Q49704;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE B1496_F2_61.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U00013; AA017144.1; -.
 DR InterPro; IPR003780; COX15_CtaA.
 DR Pfam; PF02628; COX15-CtaA; 1.
 SQ SEQUENCE 223 AA; 23338 MW; 42D301144DF0B355 CRC64;

Alignment Scores:
 Pred. No.: 6.86 Length: 223
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q49704 (1-223)

QY 516 GCCACTTGCAATCATCTCTCTGTG 490
 Db 137 AlaHisLeuHisSerSerLeuVal 145

RESULT 21

O54948 PRELIMINARY; PRT; 284 AA.
 AC O54948;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Senaphorin IV homolog (Fragment).
 GN SEMA3F.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hu L.-J., Drabkin H.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF036162; AAB88861.1; -.
 DR MGD; MGI:1096347; Sema3f.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 FT NON TER 284 284
 SQ SEQUENCE 284 AA; 32151 MW; BFD8D7E66BD08027 CRC64;

Alignment Scores:
 Pred. No.: 6.63 Length: 284
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x O54948 (1-284)

QY 1106 ACTCATTTTGATGAAGTCAGGATGTA 1132
 Db 235 ThrHisPheAspGluLeuGlnAspVal 243

RESULT 22

O33142 PRELIMINARY; PRT; 311 AA.
 ID O33142
 AC O33142;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 32.6 kDa protein (Putative membrane protein).
 DE ML0587 OR MLCL536.34.
 GN Mycobacterium leprae.
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K., Harris D.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RC MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 DR EMBL; Z99125; CAB16178.1; -.
 DR EMBL; AL583919; CAC30095.1; -.
 DR Leproma; ML0587; -.
 DR InterPro; IPR003780; COX15_CtaA.
 DR Pfam; PF02628; COX15-CtaA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 311 AA; 32590 MW; 8F5B91AAA6378E08 CRC64;

Alignment Scores:
 Pred. No.: 6.55 Length: 311
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x O33142 (1-311)

QY 516 GCCCAGTTCATTCATCTCTCTCTG 490
 Db 225 AlahieLeuHieserSerLeuLeuVal 233

RESULT 23
 Q8BW98
 ID Q8BW98 PRELIMINARY; PRT; 325 AA.
 AC Q8BW98;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Similar to eukaryotic translation initiation factor 3, subunit 8 (110kD).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC005333; AAH00533.1; -.
 DR InterPro; IPR000717; FCI.
 DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 KW Initiation factor.
 SQ SEQUENCE 325 AA; 37680 MW; 701BESA01DD7E10B CRC64;

Alignment Scores:
 Pred. No.: 6.51 Length: 325
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BW98 (1-325)

QY 2093 CAAGCCTTCGTCTAGTCTACAA 2119
 Db 60 GlnGlyLeuLeuArgSerLeuGln 68

RESULT 24
 Q8BN9
 ID Q8BN9 PRELIMINARY; PRT; 333 AA.
 AC Q8BN9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein NT2RP2004036.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
 RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
 RA Nagahari K., Sugano S., Isogai T.;
 RT "HRI human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK075388; BAC11588.1; -.
 DR InterPro; IPR003599; IG-
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin-like.
 DR Pfam; PF01437; PSI; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Hypothetical protein.

SQ SEQUENCE 333 AA; 35797 MW; F74AF2628FA274F3 CRC64;

Alignment Scores:
 Pred. No.: 6.48 Length: 333
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8NEN9 (1-333)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
 Db 11 LeuAlaArgAspProTyrCysAlaTir 19

RESULT 25
 Q8F7R6
 ID Q8F7R6 PRELIMINARY; PRT; 350 AA.
 AC Q8F7R6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Dna protein.
 GN DSHA OR LA0878.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE011272; AAN48077.1; -.
 KW Complete proteome.
 SQ SEQUENCE 350 AA; 39815 MW; 78044D56923114CD CRC64;

Alignment Scores:
 Pred. No.: 6.44 Length: 350
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8F7R6 (1-350)

QY 798 CTATCTCCGAACCTCTGGCACCACC 824
 Db 188 LeuSerSerGluLeuLeuGlyThrThr 196

RESULT 26
 Q8WVP9
 ID Q8WVP9 PRELIMINARY; PRT; 458 AA.
 AC Q8WVP9;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC017658; AAH17658.1; -.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin-repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.

```
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 458 AA; 50263 MW; B8A52233DCE848E7 CRC64;

Alignment Scores:
Pred. No.: 6.2 Length: 458
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8WVP9 (1-458)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 158 LeuAlaArgAspProTyxCysAlaTrp 166

RESULT 27
Q96FK5 PRELIMINARY; PRT; 475 AA.
AC Q96FK5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010701; AAH10701.1; -.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; PSI; 1.
DR SMART; SM00423; PSI; 1.
KW Hypothetical protein.
SQ SEQUENCE 475 AA; 52298 MW; EE90909701AE2264 CRC64;

Alignment Scores:
Pred. No.: 6.17 Length: 475
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96FK5 (1-475)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 175 LeuAlaArgAspProTyxCysAlaTrp 183

RESULT 28
Q9NX92 PRELIMINARY; PRT; 510 AA.
AC Q9NX92;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ20369.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
[1]
RN SEQUENCE FROM N.A.
RP Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000376; BAA91124.1; -.
DR InterPro; IPR003599; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 510 AA; 56147 MW; 53125912E9AD0F04 CRC64;

Alignment Scores:
Pred. No.: 6.11 Length: 510
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9NX92 (1-510)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 188 LeuAlaArgAspProTyxCysAlaTrp 196

RESULT 29
Q8BIC3 PRELIMINARY; PRT; 550 AA.
AC Q8BIC3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 4B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028896; BAC26181.1; -.
SQ SEQUENCE 550 AA; 61309 MW; 8C28E29DC4CAC958 CRC64;

Alignment Scores:
Pred. No.: 6.04 Length: 550
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BIC3 (1-550)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 524 LeuAlaArgAspProTyxCysAlaTrp 532
```



```
RESULT 30
QBUC1
ID Q8BJC1 PRELIMINARY; PRT; 799 AA.
AC Q8BJC1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 4D precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK088653; BAC40480.1; -.
FT NON_TER 1
SQ SEQUENCE 799 AA; 88814 MW; 415D3B687150A59A CRC64;

Alignment Scores:
Pred. No.: 5.74 Length: 799
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BJC1 (1-799)

QY 914 GAGATGACAAAGTACTTTTCTTC 940
Db 165 GluAspAspLysValTyrPhePhe 173

RESULT 31
QBIR6
ID Q8BIR6 PRELIMINARY; PRT; 823 AA.
AC Q8BIR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 4B
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK030331; BAC26906.1; -.
SQ SEQUENCE 823 AA; 91485 MW; E3B8C5E5CFA1CD45 CRC64;

Alignment Scores:
Pred. No.: 5.71 Length: 823
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BIR6 (1-823)
```

```
QY 1784 CTCGCCGAGACCTTACTGTGTGG 1810
Db 524 LeuAlaArgAspProTyrCysAlaTyr 532

RESULT 32
Q9C0B8
ID Q9C0B8 PRELIMINARY; PRT; 893 AA.
AC Q9C0B8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1745 (Fragment).
GN KIAA1745.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
  The complete sequences of 100 new cDNA clones from brain which code
  for large proteins in vitro."
RL DNA Res. 7:347-355 (2000).
DR EMBL; AB051532; BAB21836.1; -.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR SMART; SM00403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 893 AA; 98361 MW; D6C6C48DEE524F14 CRC64;

Alignment Scores:
Pred. No.: 5.65 Length: 893
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9C0B8 (1-893)

QY 1784 CTCGCCGAGACCTTACTGTGTGG 1810
Db 593 LeuAlaArgAspProTyrCysAlaTyr 601

RESULT 33
Q8R3M7
ID Q8R3M7 PRELIMINARY; PRT; 911 AA.
AC Q8R3M7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to eukaryotic translation initiation factor 3, subunit 8
  (110kD).
GN 3230401013RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC025032; AAH25032.1; -.
DR MGD; MGI:1913935; 3230401013RIK.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR000717; PCI.
```

DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
 KW Initiation factor.
 SQ SEQUENCE 911 AA; 105545 MW; 5A2D29E961073084 CRC64;

Alignment Scores:
 Pred. No.: 5.63 Length: 911
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8R3M7 (1-911)

QY 2093 CAAGGCTTCCTAGTACGTAGTACAA 2119
 Db 646 GlnGlyLeuLeuArgSerLeuGln 654

RESULT 34
 Q8R1B4 PRELIMINARY; PRT; 911 AA.
 ID Q8R1B4
 AC Q8R1B4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to eukaryotic translation initiation factor 3, subunit 8 (110KD).
 DE (110KD).
 GN 3230401013RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC024855; AH24855.1;
 DR MGD; MGI:1913935; 3230401013RIK.
 DR InterPro; IPR001395; Aldo/ket_red.
 DR InterPro; IPR000717; PCI.
 DR SMART; SM00088; PINT; 1.
 DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
 KW Initiation factor.
 SQ SEQUENCE 911 AA; 105531 MW; A88282C5B4C55AEB CRC64;

Alignment Scores:
 Pred. No.: 5.63 Length: 911
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8R1B4 (1-911)

QY 2093 CAAGGCTTCCTAGTACGTAGTACAA 2119
 Db 646 GlnGlyLeuLeuArgSerLeuGln 654

RESULT 35
 Q8R1B4 PRELIMINARY; PRT; 4859 AA.
 ID Q8R1B4
 AC Q8R1B4;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ryanodine receptor type 3.
 GN RYR3.
 OS Mus musculus (American mink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
 OC Mustela.
 OX NCBI_TaxID=9667;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marziani G., Rossi D., Giannini G., Charlesworth A., Sorrentino V.;
 RT "cDNA cloning reveals a tissue specific expression of alternatively
 RT spliced transcripts of the ryanodine receptor type 3 (RYR3) calcium
 RT release channel";
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RL EMBL; Y07749; CAA69029.1;
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR005797; Cytb_56_N.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RyR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF00520; Ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR Ionic channel; Receptor; Transmembrane.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4859 AA; 550457 MW; B3B5EC108C3A653D CRC64;

Alignment Scores:
 Pred. No.: 4.46 Length: 4859
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x Q95201 (1-4859)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
 Db 2132 ArgGlySerThrProLeuAspValala 2140

RESULT 36
 Q91319 PRELIMINARY; PRT; 4868 AA.
 ID Q91319
 AC Q91319;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Ryanodine receptor beta isoform.
 GN FROG-BRR.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=94274714; PubMed=8006029;
 RA Oyama H., Murayama T., Takagi T., Iino M., Iwabe N., Miyata T.,
 RA Oyama Y., Endo M.;
 RT "Primary structure and distribution of ryanodine-binding protein
 RT isoforms of the bullfrog skeletal muscle";
 RL J. Biol. Chem. 269:17206-17214(1994).
 DR EMBL; D21071; BAA04647.1;
 DR InterPro; IPR000699; Ca-rel_channel.

QY 67 CGTGGTGCACACCCCTCGACGTGGCA 41
 DB 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 39
 Q9ERN6 PRELIMINARY; PRT; 4967 AA.
 AC Q9ERN6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Cardiac Ca2+ release channel.
 GN 9330127120RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99403027; PubMed=10473538;
 RA Zhao M., Li P., Li X., Zhang L., Winkfein R.J., Chen S.R.;
 RT "Molecular identification of the ryanodine receptor pore-forming
 segment.";
 RL J. Biol. Chem. 274:25971-25974(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen S.W.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF295105; AAG34081.1; -;
 DR MGD; MGI:1924803; 9330127120RIK.
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR PRINTS; PR00795; RYANODINER.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR Ionic channel; Transmembrane.
 KW SMART; SM00449; SPRY; 3.
 SQ SEQUENCE 4967 AA; 564887 MW; 81A27F4DCFF4E4FB CRC64;

Alignment Scores:
 Pred. No.: 4.44 Length: 4967
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q9ERN6 (1-4967)

QY 67 CGTGGTGCACACCCCTCGACGTGGCA 41
 DB 2234 ArgGlySerThrProLeuAspValAla 2242

RESULT 40
 Q29621 PRELIMINARY; PRT; 4968 AA.
 AC Q29621;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Cardiac RYANODINE receptor.

GN RYR-2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart, and Brain;
 RX MEDLINE=91032060; PubMed=2226801;
 RA Nakai J., Imagawa T., Hakamat Y., Shigekawa M., Takeshima H., Numa S.;
 RT "Primary structure and functional expression from cDNA of the cardiac
 ryanodine receptor/calcium release channel.";
 RL FEBS Lett. 271:169-177(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart, and Brain;
 RX MEDLINE=93147068; PubMed=1337083;
 RA Imagawa T., Nakai J., Takeshima H., Nakasaki Y., Shigekawa M.;
 RT "Expression of Ca(2+)-induced Ca2+ release channel activity from
 cardiac ryanodine receptor cDNA in Chinese hamster ovary cells.";
 RL J. Biochem. 112:508-513(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart, and Brain;
 RA Nakai J.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U50465; AAA93465.1; -;
 DR EMBL; U50466; AAA93469.1; -;
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR Ionic channel; Transmembrane.
 KW SMART; SM00449; SPRY; 3.
 SQ SEQUENCE 4968 AA; 565081 MW; 04B16A35D06B852D CRC64;

Alignment Scores:
 Pred. No.: 4.44 Length: 4968
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x Q29621 (1-4968)

QY 67 CGTGGTGCACACCCCTCGACGTGGCA 41
 DB 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 41
 Q91313 PRELIMINARY; PRT; 5037 AA.
 AC Q91313;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Ryanodine receptor alpha isoform.
 GN FROG-ARR.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DR InterPro; IPR003608; MIR.
DR InterPro; IPR001215; RYAN.
DR InterPro; IPR003032; RYR.
DR InterPro; IPR003877; SPRY receptor.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITPR; 2.
DR Pfam; PF02026; RYR; 4.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Ionic channel; Receptor; Transmembrane.
SQ SEQUENCE 5081 AA; 576220 MW; 5A3D7E253CFBC09B CRC64;

Alignment Scores:

Pred. No.:	4.43	Length:	5081
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.03%	Indels:	0
DB:	13	Gaps:	0

US-09-774-490-1 (1-2709) x 013054 (1-5081)

QY 67 CGTGGTGCACACCCCTCGAGTGGCA 11
|||||
DB 2306 ARGGLYSerThrProLeuAspValala 2314

RESULT 43

Q24498	PRELIMINARY;	PRT; 5107 AA.
ID Q24498	PRELIMINARY;	PRT; 5107 AA.
OC Q24498	Q24499; Q24500; Q24501; Q9VAY7;	
DC 01-NOV-1996	(TrEMBLrel. 01, Created)	
DT 01-JUN-2000	(TrEMBLrel. 14, Last sequence update)	
DT 01-OCT-2002	(TrEMBLrel. 22, Last annotation update)	
DE RYanodine receptor.		
DE RYA-R44F OR CG10844.		
OS Drosophila melanogaster (Fruit fly).		
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC Ephydroidea; Drosophilidae; Drosophila.		
ON NCBI_TaxID=7227;		
OX [1]		
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4 AND 5), AND CHARACTERIZATION.		
RX MEDLINE=94102409; PubMed=8276118;		
RX Takeshima H., Nishi M., Iwabe N., Miyata T., Hosoya T., Maesai I.,		
RX Hotta Y.		
RT "Isolation and characterization of a gene for a ryanodine		
RL receptor/calcium release channel in Drosophila melanogaster.";		
RL FEBS Lett. 337:81-87(1994).		
RN [2]		
RP SEQUENCE FROM N.A. (ISOFORM 1).		
RC STRAIN=BERKELEY;		
RC MEDLINE=20196006; PubMed=10731132;		
RX Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA Ananatzides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,		
RA Sutcon G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,		
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkoch C., Baldwin D.,		
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA Beeson K.Y., Benson P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: COMMUNICATION BETWEEN TRANSVERSE-TUBULES AND
CC SARCOPLASMIC RETICULUM. CONTRACTION OF MUSCLE IS TRIGGERED BY
CC RELEASE OF CA++ FROM SR FOLLOWING DEPOLARIZATION OF T-TUBULES (BY
CC SIMILARITY).
CC -!- FUNCTION: INTRACELLULAR CALCIUM CHANNEL MAY BE INVOLVED IN
CC MORPHOGENETIC MOVEMENTS DURING EMBRYONIC DEVELOPMENT.
CC -!- SUBUNIT: HOMOTETRAMER (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4 AND 5;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DEVELOPMENTAL STAGE: ABUNDANT IN 6-12 HOUR EMBRYOS, REDUCED
CC EXPRESSION IN SECOND AND THIRD INSTAR LARVAL STAGES.
CC -!- MISCELLANEOUS: THE CALCIUM RELEASE CHANNEL IS MODULATED BY CA++,
CC MG++, ATP, AND CALMODULIN (BY SIMILARITY).
CC -!- MISCELLANEOUS: THE CALCIUM RELEASE CHANNEL ACTIVITY RESIDES IN THE
CC C-TERMINAL REGION WITH THE REMAINING PART OF THE PROTEIN
CC CONSTITUTES THE 'FOOT' STRUCTURE SPANNING THE JUNCTIONAL GAP
CC BETWEEN THE SR AND THE T-TUBULE. IT IS POSSIBLE THAT THE FOOT
CC STRUCTURE INTERACTS WITH THE CYTOPLASMIC REGION OF THE
CC DIHYDROPYRIDINE RECEPTOR (BY SIMILARITY).
CC -!- MISCELLANEOUS: RYANODINE IS AN ALKALOID THAT BINDS TO THE CA-
CC RELEASE CHANNEL IN JUNCTIONAL SR AND MODULATES ITS ACTIVITY (BY
CC SIMILARITY).
CC -!- SIMILARITY: LOCAL & LOW WITH THE NICOTINIC ACETYLCHOLINE RECEPTOR
CC (N-ACHR) SUBUNITS.
CC EMBL; D17389; BAA04212.1; ALT INIT.
CC EMBL; D17389; BAA41469.1; ALT INIT.
CC EMBL; D17389; BAA41470.1; ALT INIT.
CC EMBL; D17389; BAA41471.1; ALT INIT.
CC EMBL; AE003835; AAF59036.1; -;
CC FlyBase; FBgn011286; Rva-r44F.
CC InterPro; IPR000699; Ca-rel_channel.
CC InterPro; IPR001682; Ca/Na_pore.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR003608; MIR.
CC InterPro; IPR001215; Ryanodn_receptor.
CC InterPro; IPR003032; RYR.
CC InterPro; IPR003877; SPRY_receptor.
CC Pfam; PF00520; Ion_trans; 1.
CC Pfam; PF02815; MIR; 4.
CC Pfam; PF01365; RYDR_ITPR; 2.
CC Pfam; PF02026; RYR; 4.
CC Pfam; PF00622; SPRY; 3.
CC PRINTS; PR00795; RYANODINER.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00472; MIR; 4.
CC SMART; SM00449; SPRY; 3.
CC Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
KW Phosphorylation; Alternative splicing; Developmental protein.
FT DOMAIN 1 3207 CYTOPLASMIC.
FT TRANSMEM 3208 3228 M' (POTENTIAL).
FT TRANSMEM 3281 3299 M' (POTENTIAL).

FT TRANSMEM 4095 4114 M1 (POTENTIAL).
FT TRANSMEM 4133 4151 M2 (POTENTIAL).
FT TRANSMEM 4382 4405 M3 (POTENTIAL).
FT TRANSMEM 4447 4467 M4 (POTENTIAL).
FT TRANSMEM 4625 4646 M5 (POTENTIAL).
FT TRANSMEM 4719 4742 M6 (POTENTIAL).
FT TRANSMEM 4854 4885 M7 (POTENTIAL).
FT TRANSMEM 4902 4921 M8 (POTENTIAL).
FT TRANSMEM 4943 4962 M9 (POTENTIAL).
FT TRANSMEM 4987 5001 M10 (POTENTIAL).
FT DOMAIN 1954 1961 GLU-RICH (ACIDIC).
FT DOMAIN 825 904 6 X APPROXIMATE REPEATS.
FT REPEAT 825 936 1.
FT REPEAT 938 1055 2. (INCOMPLETE).
FT REPEAT 1382 1397 3. (INCOMPLETE).
FT REPEAT 1409 1444 4. (INCOMPLETE).
FT REPEAT 2800 2920 5.
FT REPEAT 2921 3043 6.
FT MOD_RES 2918 2918 PHOSPHORYLATION (BY CAPK AND CGPK)
FT MOD_RES 4428 4428 (BY SIMILARITY).
FT MOD_RES 1025 1059 PHOSPHORYLATION (POTENTIAL).
FT VARSPLIC 1115 1145 VRTLVYGVLDPTGEGTEALLAEQRLKFAGR -> SA
FT VARSPLIC 1856 1869 NAPGLRICGSSDGRNGTSGRTTQVRIP (IN
FT VARSPLIC 294 297 ISOFORM 2, ISOFORM 3, ISOFORM 4 AND
FT VARSPLIC 2262 2262 ISOFORM 5).
FT VARSPLIC 2530 2530 VTKMHAGSIHFVYRAGDVIGCFIDVKEO -> EEKVYG
FT VARSPLIC 2907 2907 GVSEFGKQCGPGVGVFLDLADH (IN ISOFORM 3
FT VARSPLIC 3194 3194 AND ISOFORM 4).
FT VARSPLIC 3341 3341 MISSING (IN ISOFORM 4 AND ISOFORM 5).
FT VARSPLIC 4076 4076 CLQR -> SWQ (IN REF. 1).
FT VARSPLIC 5107 AA; 578839 MW; 2AC48FDC058F36F CRC64;
SQ SEQUENCE 5107 AA; 578839 MW; 2AC48FDC058F36F CRC64;
Alignment Scores:
Pred. No.: 4.43 Length: 5107
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x Q24498 (1-5107)
Oy 67 CGTGGGTCGACACCCCTCGACGTGGCA 41
Db 2336 ArgGlySerThrProLeuAspValala 2344
RESULT 44
Q8MK55
ID Q8MK55 PRELIMINARY; PRT; 5113 AA.
AC Q8MK55
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG10844-PB.
GN RYA-R44F OR CG10844.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banson J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champs M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,
RA Pacleb J., Parasag V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003835; AAM71084.1; -;
DR FlyBase; FBgn0011286; Rva-r44F.
DR InterPro; IPR0006599; Ca-rel channel.
DR InterPro; IPR001582; Ca/Na_Pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003032; RyR.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITER; 2.
DR Pfam; PF02026; RyR; 4.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
RW Ionic channel; Transmembrane.

SQ SEQUENCE 5113 AA; 579473 MW; 6BE70740475694AB CRC64;
Alignment Scores:
Pred. No.: 4.42 Length: 5113
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x Q8MKS3 (1-5113)
QY 67 CGTGGTCCGACACCCCTCGACGTGGCA 41
DB 2342 ArgglyserthrProleuaspValala 2350
RESULT 46
Q8MKS4
ID Q8MKS4 PRELIMINARY; PRT; 5127 AA.
AC Q8MKS4;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CG10844-PC.
GN RYA-R44F OR CG10844.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y., H.C., Blazek R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Danke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman K.A., Nixon K., Nusskern D.R., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]

RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Buesam D.A.,
 RA Carlson J.W., Center A., Champ E., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrieria S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Igbegun C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mirra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003835; AAM71083.1; -;
 DR FlyBase; FBgn0011286; Rya-r44F.
 DR InterPro; IPR000699; Ca-rel channel.
 DR InterPro; IPR001682; Ca/Na pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY receptor.
 DR Pfam; PF00520; ion trans; 1.
 DR Pfam; PF02815; MIR_4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR ProDom; PD000012; EF-hand; 1.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 KW Ionic channel; transmembrane.
 SQ SEQUENCE 5127 AA; 580814 MW; C4F4F3B75DCECOAC CRC64;

Alignment Scores:
 Pred. No.: 4.42 Length: 5127
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x QBMK54 (1-5127)

Oy 67 CGTGGTCGACACCCCTCGAGTGCCA 41
 Db 2356 ArgGlySerThrProLeuAspValAla 2364

RESULT 47
 Q8K4M4

ID Q8K4M4 PRELIMINARY; PRT; 48 AA.
 AC Q8K4M4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Senaphorin 3D (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Raper J.A., Renzi M.J., Feiner L.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF164784; AAM96212.1; -;
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 FT NON TER 1 48
 FT NON TER 48 48
 SQ SEQUENCE 48 AA; 5415 MW; D2994B4F10F067C0 CRC64;
 Alignment Scores:
 Pred. No.: 108 Length: 48
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 11 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8K4M4 (1-48)
 Oy 1199 ATTTCAAGGATCAGCGTGTTG 1222
 Db 2 IlePheLySGlySerAlaValCys 9
 RESULT 48
 Q924C2
 ID Q924C2 PRELIMINARY; PRT; 51 AA.
 AC Q924C2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Post-segregation killing protein.
 GN PNDA.
 OS Shigella sonnei.
 OG Plasmid Colib-P9.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=624;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P9;
 RA Sempel G., Mizobuchi K.;
 RT "Organization and diversification of plasmid genomes: complete
 RT nucleotide sequence of the Colib-p9 genome.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021078; BAA75145.1; -;
 DR InterPro; IPR000021; Hok/gef_toxin.
 DR Pfam; PF01848; HOK_GEF; 1.
 DR PRINTS; PR00281; HOKGEFTOXIC.
 DR PROSITE; PS00556; HOK_GEF; 1.
 KW Plasmid.
 SQ SEQUENCE 51 AA; 5948 MW; 2DA97D90644FEFA9 CRC64;
 Alignment Scores:
 Pred. No.: 107 Length: 51
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q9Z4C2 (1-51)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
 DB 4 GlnArgThrPheLeuMetMetLeu 11

RESULT 49

OS9335 PRELIMINARY; PRT; 101 AA.
 AC O59335; 07, Created)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein PH1667.
 GN PH1667.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000006; BAA30779.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 101 AA; 11519 MW; 846ED3C0C2CF8E7 CRC64;

Alignment Scores:

Pred. No.: 97.1 Length: 101
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 17 Gaps: 0

US-09-774-490-1 (1-2709) x O59335 (1-101)

QY 67 CTGGTGGTGGACACCTCGACGTG 44
 DB 24 ArgGlySerThrProLeuAspVal 31

RESULT 50

Q8ZE20 PRELIMINARY; PRT; 178 AA.
 AC O8ZE20; 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DE Hypothetical protein YPO2006.
 GN YPO2006.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Hollroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 DR EMBL: AJ414151; CAC90819.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 178 AA; 19006 MW; 86757210C29B6C6C CRC64;

Alignment Scores:
 Pred. No.: 89.7 Length: 178
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8ZE20 (1-178)

QY 2517 GCTCAGACACTCTGGTGGCCCTC 2494
 DB 138 AlaGlnThrLeuLeuGlyAlaLeu 145

RESULT 51

Q9S7S7 PRELIMINARY; PRT; 180 AA.
 AC Q9S7S7; 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE FltF8.39 protein.
 GN FltF8.39 OR F8A24.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Renning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Niernman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome I BAC FltF8 genomic sequence.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC016661; AAF23308.1; --
 DR EMBL: AC015985; AAF23255.1; --
 DR InterPro; IPR000804; Clat_adaptor_s.
 DR Pfam; PF01217; Clat_adaptor_s; 1.
 KW Coat protein.
 SQ SEQUENCE 180 AA; 19727 MW; 51DB46E11C67B534 CRC64;

Alignment Scores:
 Pred. No.: 89.6 Length: 180
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 10 Gaps: 0

US-09-774-490-1 (1-2709) x Q9S7S7 (1-180)

QY 336 ATCATGTTGGATTCACACATT 313
 DB 59 IleThrLeuLeuAspSerAsnIle 66

RESULT 52

Q9MB00 PRELIMINARY; PRT; 184 AA.
 AC Q9MB00; 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Nonclathrin coat protein zeta1-COP.
 GN COP21.

```

OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Brassica.
RN NCBI_TaxID=3711;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=pekinensis;
RA Hahn Y., Chung J.H.;
RT "Identification of zeta-COP genes from various organisms.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB040039; BAA92778.1; -.
DR InterPro; IPR000504; Clat_adaptor_s.
DR Pfam; PF01217; Clat_adaptor_s; 1.
KW Coat protein.
SQ SEQUENCE 184 AA; 20344 MW; 97C83032C008E323 CRC64;

Alignment Scores:
Pred. No.: 89.3 Length: 184
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 10 Gaps: 0

US-09-774-490-1 (1-2709) x Q9MB00 (1-184)
Qy 336 ATCACATTGTTGGATTCCCAACATT 313
Db 64 ILeThrLeuLeuAspSerAsnIle 71

RESULT 53
Q8CL36 PRELIMINARY; PRT; 218 AA.
AC Q8CL36;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE Hypothetical.
GN Y2302.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=2137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611 (2002).
DR EMBL; AB013832; AAM85861.1; -.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 23531 MW; 17E162EB7B28FDB8 CRC64;

Alignment Scores:
Pred. No.: 87.2 Length: 218
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8CL36 (1-218)
Qy 2517 GCTCAGACACTCTGGTGCCCTC 2494
Db 178 AlaGlnThrLeuLeuGlyAlaLeu 185

```

```

RESULT 54
Q9KAM9 PRELIMINARY; PRT; 223 AA.
AC Q9KAM9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Butyrate-acetoacetate CoA-transferase.
GN BH2258.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AP001514; BAB05977.1; -.
DR InterPro; IPR004165; CoA_trans.
DR Pfam; PF01144; CoA_trans; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 223 AA; 24448 MW; 2CCE9F28C6FB2B26 CRC64;

Alignment Scores:
Pred. No.: 86.9 Length: 223
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9KAM9 (1-223)
Qy 2120 GTTGTAGACTAGTAGCAGAGGC 2097
Db 36 ValValAspTyrValAlaGluGly 43

RESULT 55
Q8HM61 PRELIMINARY; PRT; 227 AA.
AC Q8HM61;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE ATPase subunit 6.
GN ATPase 6.
OS Lophius americanus (American goosefish) (Anglerfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
[1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Miya M., Takeshima H., Endo H., Ishiguro N.B., Inoue J.G., Mukai T.,
RA Satoh T.P., Yamaguchi M., Kawaguchi A., Mabuchi K., Shirai S.M.,
RA Nishida M.;
RT "Major patterns of higher teleostean phylogenies: A new perspective
RT based on 100 complete mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 26:121-138 (2002).
DR EMBL; AP004414; BAC23299.1; -.
KW Mitochondrion.
SQ SEQUENCE 227 AA; 25102 MW; 4312A5A602AD4D8D CRC64;

```

Alignment Scores:
 Pred. No.: 86.7 Length: 227
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 8 Gaps: 0

US-09-774-490-1 (1-2709) x Q8HM61 (1-227)

QY 1248 CTTTCACATCACTCATGCTATAC 1225
 |||||
 Db 72 LeuLeuThrSerLeuMetLeuTyr 79

RESULT 56
 Q96ID9 PRELIMINARY; PRT; 277 AA.
 ID Q96ID9
 AC Q96ID9;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; BC007602; AA07602.1; -
 DR InterPro; IPR002350; kazal.
 DR InterPro; IPR001841; ZnF_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 277 AA; 30443 MW; EBED1BD2B7B243F0 CRC64;

Alignment Scores:
 Pred. No.: 84.3 Length: 277
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96ID9 (1-277)

QY 1801 GTAAGGTCCTCGGCGAGCAACA 1778
 |||||
 Db 204 ValArgValSerGlyGluAlaThr 211

RESULT 57
 Q9V3F8 PRELIMINARY; PRT; 280 AA.
 ID Q9V3F8
 AC Q9V3F8;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE P5CR protein.
 GN P5CR OR CG6009.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS;
 RA Misener S.R., Walker V.K.;
 RT "High density of unrelated genes showing overlapping and intratronic
 transcription units in Drosophila."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003724; AAF55626.1; -
 DR EMBL; AF170829; AAD49740.1; -
 DR FlyBase; FBgn0015781; P5CR.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR TIGRFAMs; TIGR00112; proC; 1.
 DR PROSITE; PS00521; P5CR; 1.
 DR Oxidoreductase.
 SQ SEQUENCE 280 AA; 29621 MW; D3D4FD9611574684 CRC64;

Alignment Scores:
 Pred. No.: 84.2 Length: 280
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x Q9V3F8 (1-280)

QY 2520 GCAGCTCAGACACTCTGGGTGCC 2497
 |||||
 Db 215 AlaAlaGlnThrLeuLeuGlyAla 222

RESULT 58
 O96643

ID O96643 PRELIMINARY; PRT; 280 AA.
AC Q96643;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Pyroline 5-carboxylate reductase.
GN P5CR OR C6609.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS;
RA Misener S.R., Walker V.K.;
RT "Cloning and sequence analysis of the gene encoding pyroline 5-
RT carboxylate reductase from Drosophila melanogaster.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF098020; AAC70780.1; -.
DR FlyBase; FBgn0015781; P5CR.
DR InterPro; IPR00304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR TIGRFAMs; TIGR00112; proC; 1.
DR PROSITE; PS00521; P5CR; 1.
SQ SEQUENCE 280 AA; 29688 MW; 2A7416D39D2A2369 CRC64;
Alignment Scores:
Pred. No.: 84.2 Length: 280
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x O96643 (1-280)
QY 2520 GCAGCTCAGACACTCTGGTGCC 2497
Db 215 AlaAlaGlnThrLeuLeuGlyAla 222
RESULT 59
O22841
ID O22841 PRELIMINARY; PRT; 283 AA.
AC O22841;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative endochitinase.
GN AT2G43620.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.B., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrara A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RN Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002333; AAB64044.1; -.
DR HSP; P02876; 9WGA.
DR InterPro; IPR001002; Chitin binding_1.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR PRINTS; PR00451; CHITINBINDING.
DR PRODom; PD000609; Chitin binding_1; 1.
DR SMART; SM00270; ChtBD1; 1.
DR PROSITE; PS00026; CHITIN_BINDING; 1.
KW Chitin-binding.
SQ SEQUENCE 283 AA; 30377 MW; D758222971EB8404 CRC64;
Alignment Scores:
Pred. No.: 84.1 Length: 283
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 10 Gaps: 0
US-09-774-490-1 (1-2709) x O22841 (1-283)
QY 2319 TGCAGCTCATCAACACCCCAATC 2342
Db 64 CysSerSerThrThrProile 71
RESULT 60
Q91632
ID Q91632 PRELIMINARY; PRT; 292 AA.
AC Q91632;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE BTBB.
GN GENE 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U35408; AAC59863.1; -.
DR HSP; P08047; LSP2
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf_C2H2; 3.
DR PRODom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Zinc; Zinc-finger.
FT VARIANT 16 16 V -> M.
SQ SEQUENCE 292 AA; 32789 MW; EF62BD68DD9CA2C3 CRC64;
Alignment Scores:
Pred. No.: 83.7 Length: 292
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 13 Gaps: 0
US-09-774-490-1 (1-2709) x Q91632 (1-292)
QY 2316 TCATGCAGCTCATCAACACCCCA 2339

```
|||||
Db 140 SerCysSerSerThrThrPro 147
RESULT 61
Q91633 PRELIMINARY; PRT; 292 AA.
AC Q91633;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BTEB.
GN GENE 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U35409; AAC59864.1; -
DR HSSP; P08047; 1SP2.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
DR Metal-binding; Zinc; Zinc-finger.
KW SEQUENCE 292 AA; 32748 MW; 812BA9D57EF8CFA CRC64;

Alignment Scores:
Pred. No.: 83.7 Length: 292
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q91633 (1-292)

QY 2316 TCATCGAGCTCATCAACACCCCA 2339
Db 140 SerCysSerSerThrThrPro 147
RESULT 62
Q9D3M7 PRELIMINARY; PRT; 294 AA.
AC Q9D3M7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Arac-type DNA-binding domain-containing protein.
GN V21661.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016813; AAC08519.1; -
DR DNA-binding; Complete proteome.
KW SEQUENCE 294 AA; 33902 MW; A62C99A081BE20E2 CRC64;
```

```
Alignment Scores:
Pred. No.: 83.6 Length: 294
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9D3M7 (1-294)

QY 991 TCTAGCGTGTAGCTTTTCCAGA 968
Db 91 SerSerValSerSerPheSerArg 98
RESULT 63
Q9D3Q3 PRELIMINARY; PRT; 299 AA.
AC Q9D3Q3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Spermidine synthase.
GN V21635.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016813; AAC08493.1; -
DR Complete proteome.
KW SEQUENCE 299 AA; 34010 MW; 31687212D98C7C6A CRC64;

Alignment Scores:
Pred. No.: 83.4 Length: 299
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9D3Q3 (1-299)

QY 9 ATTTATCGATGTTAACAAGCTTA 32
Db 237 IleLeuSerMetLeuThrSerLeu 244
RESULT 64
Q9A6S1 PRELIMINARY; PRT; 309 AA.
AC Q9A6S1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein CC2012.
GN CC2012.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
```


RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of *Caulobacter crescentus*,"
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005874; AAK23987.1; -.
DR TIGR; CC2012; -.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR002965; P_Rich_extensn.
DR PRINTS; PR01217; PRICHEXTENSN.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 309 AA; 34342 MW; 73FC082AE5D856A8 CRC64;

Alignment Scores:
Pred. No.: 83 Length: 309
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9A651 (1-309)

Qy 1074 ATCAGACGAGCTTTGAGGAATGTT 1051
|||||
Db 38 ILeArgAlaLeuArgAsnVal 45

RESULT 65

ID Q8NH68 PRELIMINARY; PRT; 312 AA.
AC Q8NH68;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Seven transmembrane helix receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tautsuni S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
receptor genes";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB065521; BAC05769.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; Transmembrane.
SQ SEQUENCE 312 AA; 34897 MW; 851854030D2BEE6A CRC64;

Alignment Scores:
Pred. No.: 82.9 Length: 312
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8NH68 (1-312)

Qy 1770 GTGCTGAGTGTGCTCGCCGAG 1793
|||||
Db 225 ValLeuSerValAlaSerProGlu 232

RESULT 66

Q97D40

ID Q97D40 PRELIMINARY; PRT; 326 AA.
AC Q97D40;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Oligopeptide ABC transporter, ATPase component.
GN CAC3641.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium *Clostridium acetobutylicum*,"
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007859; AAK81563.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Complete proteome.
SQ SEQUENCE 326 AA; 36325 MW; 9F1A2DE19ACAD435 CRC64;

Alignment Scores:
Pred. No.: 82.4 Length: 326
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q97D40 (1-326)

Qy 296 AAATTATCTACAAAGAAATGTTG 319
|||||
Db 84 LysLeuSerTyrIysGluMetLeu 91

RESULT 67

ID Q91914 PRELIMINARY; PRT; 327 AA.
AC Q91914;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Homeobox protein.
GN XOM.
OS *Xenopus laevis* (African clawed frog);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; *Xenopus*.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96312918; PubMed=8756284;
RA Lader R.K., Mohun T., Smith C., Snape A.M.;
RT "Xom: a *Xenopus* homeobox gene which mediates the early effects of BMP-
4,"
RT Development 122:2385-2394 (1996).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; X98454; CAA67093.1; -.
DR HSSP; P02836; IENH.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.

DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS0071; HOMEBOX_2; 1.
 DR DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 327 AA; 36675 MW; AB669084F355F4E4 CRC64;

Alignment Scores:
 Pred. No.: 82.4 Length: 327
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q91914 (1-327)

QY 2310 GAGACTTCATGCAGCTCATCAACC 2333

Db 145 GluThrSerCysSerSerSerThr 152

RESULT 68

Q97D45 PRELIMINARY; PRT; 350 AA.
 AC Q97D45;

DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Oligopeptide ABC transporter, ATPase component.
 GN CAC3635.

OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J.; Breton G.; Omelchenko M.V.; Makarova K.S.; Zeng Q.,
 RA Gibson R.; Lee H.M.; Dubois J.; Qiu D.; Hitti J.; Wolf Y.I.,
 RA Tatusov R.L.; Sabathe F.; Doucette-Stamm L.; Soucaille P.; Daly M.J.,
 RA Bennett G.N.; Koonin E.V.; Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).

DR EMBL; AE007859; AAK81558.1; -.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00005; ABC tran; 1.
 DR ProDom; PD000006; ABC transporter; 1.
 DR SMART; SM00382; AAA_1
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Complete proteome.
 SQ SEQUENCE 350 AA; 39445 MW; 338619DF480EB51D CRC64;

Alignment Scores:
 Pred. No.: 81.6 Length: 350
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q97D45 (1-350)

QY 286 AAATATCTCAAGAAGATGTTG 319

Db 108 LysLeuSerTyrLysGluMetLeu 115

RESULT 69

Q9BYE7 PRELIMINARY; PRT; 352 AA.
 ID Q9BYE7
 AC Q9BYE7;

DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE HMBLR protein.
 GN HMBLR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Takahashi N.; Akasaka T.; Nakayama T.; Koseki H.; Koga H.;
 RT "Identification and characterization of hMBLR.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL; AB047006; BAB40779.1; -.
 DR EMBL; BC010235; AAH10235.1; -.
 DR InterPro; IPR002350; Kazal.
 DR InterPro; IPR001841; Znf ring.
 DR Pfam; PF00097; zf-C3HC4_1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; zinc-finger.
 SQ SEQUENCE 352 AA; 39241 MW; 559F971A86D9A093 CRC64;

Alignment Scores:
 Pred. No.: 81.5 Length: 352
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9BYE7 (1-352)

QY 1801 GTAAGGCTCTCGGCGAGCAACA 1778

Db 279 ValArgValSerGlyGluAlaThr 286

RESULT 70

Q99NA9 PRELIMINARY; PRT; 353 AA.
 ID Q99NA9
 AC Q99NA9;

DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE HMBLR protein (RIKEN CDNA 4933407A11 gene).
 GN 4933407A11RIK OR HMBLR.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RA Takahashi N.; Akasaka T.; Nakayama T.; Koseki H.; Koga H.;
 RT "Identification and characterization of hMBLR.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL; AB047007; BAB40780.1; -.
 DR EMBL; BC016195; AAH16195.1; -.

DR MGI:1918291; 4933407A11RIK.
 DR InterPro; IPR002350; kazal.
 DR InterPro; IPR001841; Znf ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 DR PROSITE; PS00518; ZF_RING_2; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 353 AA; 39820 MW; OD29B7FE065B5448 CRC64;

Alignment Scores:
 Pred. No.: 81.5 Length: 353
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q99NA9 (1-353)

QY 1801 CTAGGGTCTCGGCGAGGCAACA 1778
 |||||
 Db 280 ValArgValSerGlyGluAlaThr 287

RESULT 71
 Q90753
 ID Q90753 PRELIMINARY; PRT; 371 AA.
 AC AC Q90753;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Host shut-off factor (Fragment).
 GN UL41.
 OS Feline herpesvirus (Feline herpesvirus 1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B927;
 EX MEDLINE=98141690; PubMed=9482586;
 RA Willoughby K., Bennett M., Williams R.A., McCracken C., Gaskell R.M.;
 RT "Sequences of the ribonucleotide reductase-encoding genes of feline
 herpesvirus 1 and molecular phylogenetic analysis.";
 RL Virus Genes 15:203-218 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B927;
 RA Willoughby K.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ006454; CAA07026.1; -;
 DR InterPro; IPR000513; EXO_N_I.
 DR InterPro; IPR002327; Virus_HS.
 DR Pfam; PF01550; Virus_HS; 1.
 FT NON_TER 1
 SQ SEQUENCE 371 AA; 42219 MW; 10672R8EFDCAP675 CRC64;

Alignment Scores:
 Pred. No.: 81 Length: 371
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 12 Gaps: 0

US-09-774-490-1 (1-2709) x Q90753 (1-371)

QY 2498 CCTCTCAAAATCGTGGTCTCTCC 2475
 |||||
 Db 289 ProSerGlnIleArgGlySerSer 296

RESULT 72

Q9AQP4
 ID Q9AQP4 PRELIMINARY; PRT; 386 AA.
 AC Q9AQP4;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Putative ABC transporter subunit.
 OS Pseudomonas sp. CA10.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=135214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 EX MEDLINE=21264379; PubMed=11371531;
 RA Nojiri H., Sekiguchi H., Maeda K., Urata M., Nakai S., Yoshida T.,
 RA Habe H., Omori T.;
 RT "Genetic characterization and evolutionary implications of car gene
 cluster in carbazole-degrader, Pseudomonas sp. strain CA10.";
 RL J. Bacteriol. 183:3663-3679 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 EX MEDLINE=97386424; PubMed=9244273;
 RA Sato S., Ouchiya N., Kimura T., Nojiri H., Yamane H., Omori T.;
 RT "Cloning of genes involved in carbazole degradation of Pseudomonas sp.
 strain CA10: nucleotide sequence of genes and characterization of
 RT meta-cleavage enzymes and hydrolase.";
 RL J. Bacteriol. 179:4841-4849 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 EX MEDLINE=97386425; PubMed=9244274;
 RA Sato S., Nam J., Kasuga K., Nojiri H., Yamane H., Omori T.;
 RT "Identification and characterization of genes encoding carbazole 1,9a-
 RT dioxygenase in Pseudomonas sp. strain CA10.";
 RL J. Bacteriol. 179:4850-4858 (1997).
 DR EMBL; AB047548; BAB32743.1; -;
 SQ SEQUENCE 386 AA; 41689 MW; 7B6F4D2292D46BCE CRC64;

Alignment Scores:
 Pred. No.: 80.5 Length: 386
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q9AQP4 (1-386)

QY 1734 TCCCTTACACCGGTGTGATTT 1757
 |||||
 Db 92 SerLeuTyThrGlyValIlePhe 99

RESULT 73
 Q8GI39
 ID Q8GI39 PRELIMINARY; PRT; 386 AA.
 AC Q8GI39;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative ABC transporter subunit.
 OS Pseudomonas resinovorans.
 OG Plasmid pCARI.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=53412;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Maeda K., Nojiri H., Shintani M., Yoshida T., Habe H., Omori T.;
 RT "Complete nucleotide sequence of carbazole/dioxin-degrading plasmid
 pCARI in Pseudomonas resinovorans strain CA10 indicates its mosaicity
 RT and the presence of large catabolic transposon Tn4676.";

RL J. Mol. Biol. 0:0-0(2002).
DR EMBL; AB088420; BAC41522.1; --
KW Plasmid.
SQ SEQUENCE 386 AA; 41461 MW; C9A34BAF2DD24F7A CRC64;

Alignment Scores:
Pred. No.: 80.5 Length: 386
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q8GI39 (1-386)

QY 1734 TCCTTTACACGGTGTGATATT 1757

Db 92 SerLeuTyThrGlyValIlePhe 99

RESULT 74

Q8RTH6 PRELIMINARY; PRT; 402 AA.
AC Q8RTH6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative O antigen ligase.
GN WAAL.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]_TaxID=666;
RP SEQUENCE FROM N.A.
RA Nesper J., Kraiss A., Schild S., Blass J., Klose K.E., Bockemuehl J.,
RA Reidl J.;
RT "Comparative and genetic analysis of the putative Vibrio cholerae LPS
core oligosaccharide biosynthesis (waa) gene cluster";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF49194; AAL77364.1; --
DR InterPro; IPR000504; RNA_rec_mot.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Ligase.
SQ SEQUENCE 402 AA; 46039 MW; B44948E5DF833C0E CRC64;

Alignment Scores:
Pred. No.: 80 Length: 402
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q8RTH6 (1-402)

QY 761 AGTATAATCTCCATCTATATAA 738

Db 347 SerIleIleLeuHisLeuLeuLys 354

RESULT 75

Q52274 PRELIMINARY; PRT; 409 AA.
AC Q52274;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Specificity subunit.
GN WSDS.
OS Lactococcus lactis subsp. lactis bv. diacetylactis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=44688;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=UKJ9161;
RX MEDLINE=20541281; PubMed=11092734;
RA Deng Y.W., Liu C.Q., Dunn N.W.;
RT "Lidi, a plasmid-encoded type I restriction and modification system in
RT Lactococcus lactis";
RL DNA Seg. 11:239-245(2000).
DR EMBL; AF034786; AAB91417.1; --
DR InterPro; IPR000055; Rest_mod_DNA.
DR Pfam; PF01420; Methylase_S; 2.
KW Plasmid.
SQ SEQUENCE 409 AA; 46727 MW; 3B26DDA650A274BC CRC64;

Alignment Scores:
Pred. No.: 79.9 Length: 409
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x O52274 (1-409)

QY 1135 GAATACATCTCGAGTTCATCAAA 1112

Db 117 GluTyriLeuGlnPheIleLys 124

Search completed: August 3, 2003, 10:56:00
Job time : 268.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:37 ; Search time 86.5 Seconds
(without alignments)
9941.972 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 atctttttatttcgatg.....aggcttttttctcctaataacc 2709

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2062474

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters: -MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO spool/US09774490/runat 03082003.102927.4441/app query.fasta.1.2887
-DB=A_Geneseq_19Jun03 -QFWT=fascan -SUFFIX=oligo.rag -MINTRANS=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490@cgn 1.1.114 @runat 03082003.102927.4441 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	771	86.1	771	16	AAR71380	Human semaphorin I
2	771	86.1	771	22	AAG62726	Amino acid sequenc
3	771	86.1	771	23	ABG96413	Human ovarian canc
4	437	48.8	796	19	AAV21264	Human semaphorin I
5	434	48.5	477	16	AA874175	Human collapsin.
6	151	16.9	151	22	ABG54800	Human liver peptid
7	151	16.9	151	23	ABG42930	Human peptide enco
8	81	9.1	81	19	AAV21293	Human semaphorin I
9	77	8.6	77	19	AAV21288	Human semaphorin I
10	74	8.3	74	19	AAV21291	Human semaphorin I
11	69	7.7	69	19	AAV21274	Human semaphorin I
12	64	7.2	64	19	AAV21272	Human semaphorin I
13	57	6.4	57	19	AAV21299	Human semaphorin I
14	55	6.1	85	19	AAV21271	Human semaphorin I
15	53	5.9	57	22	ABG54764	Human liver peptid
16	53	5.9	57	22	ABG39696	Peptide #7202 enco
17	53	5.9	57	22	AAW60414	Human brain expres
18	53	5.9	57	22	AAW73050	Human bone marrow
19	53	5.9	57	22	AAW33272	Peptide #7309 enco
20	53	5.9	57	23	ABG42894	Human peptide enco
21	51	5.7	51	19	AAV21280	Human semaphorin I
22	50	5.6	50	19	AAV21346	Human semaphorin I
23	48	5.4	48	19	AAV21324	Human semaphorin I
24	46	5.1	46	19	AAV21331	Human semaphorin I
25	42	4.7	42	19	AAV21322	Human semaphorin I
26	41	4.6	41	22	ABG55921	Human liver peptid
27	41	4.6	41	22	AAW61385	Human brain expres
28	41	4.6	41	23	ABG44063	Human peptide enco
29	33	3.7	33	19	AAV21266	Human semaphorin I
30	33	3.7	33	22	ABG54883	Human liver peptid
31	33	3.7	33	22	ABG39792	Peptide #7298 enco
32	33	3.7	33	22	ABB24410	Protein #6409 enco
33	33	3.7	33	22	AAW60513	Human brain expres
34	33	3.7	33	22	AAW73166	Human bone marrow
35	33	3.7	33	22	AAW33379	Peptide #7416 enco
36	33	3.7	33	23	ABG43013	Human peptide enco
37	31	3.5	31	19	AAV21292	Human semaphorin I
38	30	3.4	30	19	AAV21279	Human semaphorin I
39	27	3.0	27	19	AAV21269	Human semaphorin I
40	26	2.9	26	19	AAV21296	Human semaphorin I
41	25	2.8	25	19	AAV21328	Human semaphorin I
42	24	2.7	24	19	AAV21285	Human semaphorin I
43	23	2.6	23	19	AAV21337	Human semaphorin I
44	23	2.6	23	19	AAV21310	Human semaphorin I
45	23	2.6	23	19	AAV21290	Human semaphorin I
46	22	2.5	22	19	AAV21334	Human semaphorin I
47	22	2.5	67	21	AAW02551	Human secreted pro
48	21	2.3	21	19	AAV21342	Human semaphorin I
49	21	2.3	21	19	AAV21343	Human semaphorin I
50	21	2.3	21	19	AAV21278	Human semaphorin I
51	20	2.2	20	19	AAV21329	Human semaphorin I
52	19	2.1	19	19	AAV21318	Human semaphorin I
53	19	2.1	19	19	AAV21304	Human semaphorin I
54	18	2.0	18	19	AAV21309	Human semaphorin I
55	17	1.9	17	19	AAV21301	Human semaphorin I
56	17	1.9	17	22	AAW62702	Hanatoxin-like seq
57	17	1.9	779	22	AAW84219	Amino acid sequenc
58	17	1.9	779	23	AAW78481	Human ZSMF-16 Ho
59	17	1.9	782	21	AAW23609	Human secreted pro
60	17	1.9	782	22	AAW65620	Novel human protei
61	17	1.9	785	21	AAW23636	Human secreted pro
62	17	1.9	875	22	AAW65619	Novel human protei
63	16	1.8	17	22	AAW62704	Hanatoxin-like seq
64	15	1.7	15	19	AAV21300	Human semaphorin I
65	15	1.7	15	19	AAV21281	Human semaphorin I
66	15	1.7	749	22	AAW62727	Amino acid sequenc
67	15	1.7	751	20	AAW30617	Human semaphorin E

68 15 1.7 751 21 AAB28379 Clone BR533.4. Ho
69 15 1.7 751 22 AAG62728 Amino acid sequenc
70 15 1.7 751 23 ABP68623 Human pancreatic c
71 14 1.6 14 19 AAY21315 Human semaphorin I
72 14 1.6 14 19 AAY21276 Human semaphorin I
73 14 1.6 777 20 AAY27127 Human brain tissue
74 14 1.6 777 21 AAY99427 Human PRO1491 (UNQ
75 14 1.6 777 22 AAY92197 Human PRO polypept
76 14 1.6 777 22 AAB66176 Protein of the inv
77 14 1.6 777 24 ABU71285 Human PRO1491 prot
78 14 1.6 777 24 ABU65742 Human secreted/tra
79 14 1.6 777 24 ABU66075 Novel human secret
80 14 1.6 777 24 ABU67579 Human secreted/tra
81 14 1.6 777 24 ABU65437 Human PRO polypept
82 14 1.6 777 24 ABU58573 Human PRO polypept
83 14 1.6 777 24 ABU58109 Human secreted/tra
84 14 1.6 777 24 ABU57104 Human PRO polypept
85 14 1.6 777 24 ABU10683 Human secreted/tra
86 14 1.6 807 22 AAG62729 Amino acid sequenc
87 13 1.5 13 19 AAY21302 Human semaphorin I
88 13 1.5 17 19 AAY21267 Human semaphorin I
89 13 1.5 17 22 AAG62705 Hanatoxin-like seq
90 13 1.5 775 19 AAW63748 Human semaphorin.
91 13 1.5 775 20 AAY43090 Mouse semaphorin H
92 13 1.5 775 24 ABR47588 Breast cancer asso
93 13 1.5 777 20 AAY43091 Mouse semaphorin H
94 13 1.5 777 22 AAG62730 Amino acid sequenc
95 13 1.5 785 22 AAG62731 Amino acid sequenc
96 12 1.3 12 19 AAY21325 Human semaphorin I
97 12 1.3 12 19 AAY21336 Human semaphorin I
98 12 1.3 12 19 AAY21320 Human semaphorin I
99 12 1.3 12 19 AAY21270 Human semaphorin I
100 12 1.3 12 19 AAY21275 Human semaphorin I

ALIGNMENTS

RESULT 1
ID AAR71380
XX AAR71380 standard; Protein; 771 AA.
AC AAR71380;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
DE Human semaphorin III protein.
XX
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW variola major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
XX
OS Homo sapiens.
XX.
PN WO9507706-A1.
XX
PD 23-MAR-1995.
XX
PF 13-SEP-1994; 94WO-US10151.
XX
PR 13-SEP-1993; 93US-0121713.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
XX
XX WPI; 1995-131177/17.
DR N-PSDB; AAO87442.
XX
PT New class of semaphorin peptide(s) and polypeptide(s) - are
potent modulators of nerve cell growth and regeneration

XX Example 2; Page 60-63; 101pp; English.
PS The sequence of the human semaphorin III protein. The proteins
XX encoded by the grasshopper semaphorin I (AAQ87441), human semaphorin
CC III, vaccinia virus semaphorin IV (AAQ87443), Drosophila semaphorin I and
CC II (AAQ87444-5), Tribolium semaphorin I (AAQ87446) or variola major
CC (smallpox) virus semaphorin IV (AAQ87447) genes were used to generate a
CC series of peptides (AAR70370-R70418), which retain semaphorin receptor
CC binding activity. The semaphorin derived or semaphorin receptor derived
CC peptides are potent modulators of nerve cell growth, immune
CC responsiveness and viral pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
XX (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 771 AA;

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR71380 (1-771)

Qy 200 ATGGCGTGTAACTAGAGATTGCTGCTTTTGGGGAGTATTACTTACAGCAGAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
Qy 260 AACTATCAGATGCGAAGAACATGTCACAGGCTGAATATATCTCTACAAAGAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
Qy 320 GAATCCAAACATGTGATCAGTCTTTCAATGGCTGGCCCAACAGCTCCAGTATTATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
Qy 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTy-ValGlyAlaLysAspHisIlePheSerPhe 80
Qy 440 GACCTGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
Qy 500 GATGAATGCAAGTGGGCTGGAAAGACATCTGAAAGAAATGTGCTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
Qy 560 CTTAAGGCATATAATCAGACTCACTTGTACGCCCTGTGGAGCGGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProile 140
Qy 620 TGCACCTACATTGAATTGGACATCATCTCCAGGACACATATTTTAACTGGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
Qy 680 CATTGTAACACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGTGCAGAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
Qy 740 TTAATAGATGGAGAATTATATCTTGGAACTGCAGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleaspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCGGAACCTCTGGGACACCCACCACCAATCAGGACAGCAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

Qy	860	CTCAATGATCCAAAGTTCTATTAGTGGCCCACTCATCTACAGAGAGTGACAATCTCTGAAGAT	919
Db	221	LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp	240
Qy	920	GACAAAGTATACTTTTCTTCCTCGTGAATGCCAATAGATGAGGAACTCTGGAAAGCT	979
Db	241	AspLysValTyrPhePhePheArgGluAsnAlaIleaspGlyGluHisSerGlyLysAla	260
Qy	980	ACTCACGCTAGAATPAGTTCAGATATGCAAGAATGACTTTCGAGGGCACAGAAGTCTGGTG	1039
Db	261	ThrHisalaargIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal	280
Qy	1040	AATAAATGGACAACATTCCTCAAAGTCGTCGTGATTTGCTCAGTGCAGGTCCAAATGGC	1099
Db	281	AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly	300
Qy	1100	ATTGACACTCATTTTGATGAACCTGCAGGATGTATCTCTAATGAACTTTAAAGATCCTAAA	1159
Db	301	IleaspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys	320
Qy	1160	AATCCAGCTGTATATGAGTGTTTTACCACCTTCAGTAACTATTTCAAGGATCAGCCGTG	1219
Db	321	AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal	340
Qy	1220	TGTATGTATAGCATGAGTGTGAGAAGGGTGTCTCTGGTCCATATGCCCAAGGGAT	1279
Db	341	CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp	360
Qy	1280	GGACCCAACTCATNTGGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGA	1339
Db	361	GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr	380
Qy	1340	TGTCGCCAGCAAAACATTTGGTGGTGGTGGTACTCTCAAAAGAGACCTTCCTGATGATGTTATA	1399
Db	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
Qy	1400	ACCTTTGCCAAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCGCCCA	1459
Db	401	ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro	420
Qy	1460	ATAGTGATCAAAACGGATGTAAATTTACAAATTTACAAATTTGCTGTAGACCGAGTGGAT	1519
Db	421	IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp	440
Qy	1520	GCAGAAGATCGACAGTATGATGTTATGTTATCGAACACAGATGTGGACCGTTCCTTAAA	1579
Db	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
Qy	1580	GTAGTTTCAATTCCTAAGGAGACTGTGTATGATTAGAAAGAGTTCCTGCTGGAAGAAATG	1639
Db	461	ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet	480
Qy	1640	ACAGTTTTTCGGNAACCGACTGCTATTTCACGAATGGAGCTTTCACATAGCAGCAACAA	1699
Db	481	ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
Qy	1700	CTATATATTGGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGTCTGATATTTC	1759
Db	501	LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr	520
Qy	1760	GGGAAACGGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTCTGGGATGGTTCT	1819
Db	521	GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTTrpAspGlySer	540
Qy	1820	GCATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAACAAGCAAGATATAAGAAAT	1879
Db	541	AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn	560
Qy	1880	GGAGACCCACTGACTCTGTCAGACTTACCACTATCAATCAATCAATCAATCAATCAAT	1939
Db	561	GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro	580
Qy	1940	GAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTACCAATTTTGGAAATGCAGTCCGAAG	1999

Db	581	GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys	600
QY	2000	TCGCAGAGAGCGCTGGTCTATTGGCAATTCACAGAGCGAAATGAAGAGCGCAAGAAGAGAG	2059
Db	601	SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluAArgLysGluGlu	620
QY	2060	ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGTCTACGTAGTCTACAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln	640
QY	2120	CAGAAGATTACGGCAATTAACCTCTGCCATCGCGGTGGAAATGCGGTTCATACAACTCTT	2179
Db	641	GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu	660
QY	2180	CTTAAGGTAAACCCGGAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTCATAAAGAT	2239
Db	661	LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp	680
QY	2240	GATCATGGAGATGGCTCTAAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	681	AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys	700
QY	2300	GTCGTGGTACAGAGACTTCATCGAGCTCATCAACACCCCAATCTCAACAGATGGATGAG	2359
Db	701	ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
QY	2360	TTCTGTGAACAAAGTTTCGAAAGGGCCGAAACAAACGTCGGCAAGGCCAGGACATACC	2419
Db	721	PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr	740
QY	2420	CCAGGAAACAGTAAACAAATGGAAGCACTTACAAGAAAATAAGAAAGGTAGAAACAGGAGG	2479
Db	741	ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg	760
QY	2480	ACCACCAATTTGAGCGGCCACCCAGAGTGTCT	2512
Db	761	ThrHisGluPheGluArgAlaProArgSerVal	771
RESULT 2			
AAG62726			
ID	AAG62726 standard; peptide; 771 AA.		
XX			
AC	AAG62726;		
XX			
DT	17-SEP-2001 (first entry)		
XX			
DE	Amino acid sequence of human semaphorin Sema3A.		
XX			
KW	Hanatoxin; tarantula; toxin; voltage-gated potassium channel;		
KW	voltage-gated calcium channel; hanatoxin-like sequence; HTLS;		
KW	semaphorin; dorsal root ganglion repulsion; growth cone collapse.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200138491-A2.		
XX			
PD	31-MAY-2001.		
XX			
PF	07-NOV-2000; 2000WO-US41943.		
XX			
PR	08-NOV-1999; 99US-0164056.		
XX			
PA	(GEHO) GEN HOSPITAL CORP.		
XX			
PI	Behar O, Woolf CJ;		
XX			
DR	WPI; 2001-451494/48.		
XX			
PT	Polypeptide sequences that encompass the hanatoxin-like sequences of		
PT	semaphorins, useful as drugs to treat any condition or disease that		
XX	is characterized by abnormal calcium channel function -		

Polypeptide sequences that encompass the hanatoxin-like sequences of semaphorins, useful as a drugs to treat any condition or disease that is characterized by abnormal calcium channel function -

PS Claim 6; Page 11; 29pp; English.

XX The present sequence represents a semaphorin. The specification
 CC describes hantoxin-like sequences (HTLS) found in the semaphorin
 CC domain of mammalian secreted semaphorins. Hantoxin is a tarantula
 CC toxin that selectively blocks some voltage-gated potassium and calcium
 CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
 CC and growth cone collapse activities associated with semaphorins.
 CC Polypeptides containing HTLS can be used to modulate the activity of
 CC calcium channels. The peptides can also be used as an antigen to
 CC generate antibodies that can then be used to modulate the activity of
 CC calcium channels by inactivating naturally occurring channel ligands.
 CC The peptides or antibodies can be used as drugs to treat any condition
 CC or disease that is characterized by abnormal calcium channel function.

XX Sequence 771 AA;

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAG62726 (1-771)

QY	200	ATGGGCTGGTTAACTAGGATGCTGTCTTTCTGGGAGTATTACTTACAGCAAGACA	259
DB	1	MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla	20
QY	260	AACTATCAGAAATGGGAAGCAATGTGCCAGGCTGAAATATCTCAAGAAAGATGTTG	319
DB	21	AsnTyrGlnAsnGlyIleAsnAsnValProArgLeuLeuLeuSerTyrLysGluMetLeu	40
QY	320	GAATCCAAATGTGATCACTTCAATGCTTGGCCACAGCTCCAGTTATCATCCTTC	379
DB	41	GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe	60
QY	380	CTTTTGGATGAGGACGAGTAGGCTGTATGTTGGACGAAGGATCATATTTTCATTC	439
DB	61	LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe	80
QY	440	GACCTGGTTAATATCAAGATTTTCAAAGATTGTGCGCCAGTATCTTACACAGAGA	499
DB	81	AspLeuValAsnIleLysAspPheGlnLysIleValTrpValSerTyrThrArgArg	100
QY	500	GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA	559
DB	101	AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal	120
QY	560	CTTAGGCGATATATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT	619
DB	121	LeuLysAlaLysAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle	140
QY	620	TGCACCTACATTGAAATTTGGACATCATCTCAGGACAAATATTTTAAAGCTGGAGACTCA	679
DB	141	CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer	160
QY	680	CATTTTGAACCGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT	739
DB	161	HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu	180
QY	740	TTATAGATGAGATTTACTCTGGAATCGACTGATTTTATGGGCGAGACTTTGCT	799
DB	181	LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla	200
QY	800	ATCTTCCAACTCTGGGCGACACCAATCAGGACAGAGCAGCATGATTCCAGGTGG	859
DB	201	IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp	220
QY	860	CTCAATGATCAAAAGTTTCATTAGTGCCCACTCATCTCAGAGAGTGACAATCTCGAAGAT	919

DB	221	LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp	240
QY	920	GACAAAGCTATATCTTTTCTCCGTGAAATGCAATAGATGAGACATCTCTGGAAGCT	979
DB	241	AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla	260
QY	980	ACTCAGCGTAGAATAGGTACATATGCAAGAAATGACTTTTGAGGGCCACAGAGTCTGGTG	1039
DB	261	ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal	280
QY	1040	AATAAATGGACAACATCTCTCAAGCTGCTGTGATTTGCTCAGTGCCAGGTCCTCAATGGC	1099
DB	281	AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly	300
QY	1100	ATTGACACTCATTTTGTGAACTGCAGGATGTATCTCTTAATGAATTTTAAAGATCTTAAA	1159
DB	301	IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys	320
QY	1160	AATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG	1219
DB	321	AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal	340
QY	1220	TGTATGTATGATGATGATGTGAGAGGCTGCTCTTGGTCCATATGCCACAGGAT	1279
DB	341	CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp	360
QY	1280	GGACCCAACTATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCAGGCCAGGACT	1339
DB	361	GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr	380
QY	1340	TGTCCCAAGCAAAACATTTGGTGGTTTGTACTCTCAAGAGACCTTCTCTGATGATGTATA	1399
DB	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
QY	1400	ACCTTTTCAAGAAAGTATCCAGCCATCTACAATCCAGTGTTCCTATGAACAATCGCCCA	1459
DB	401	ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro	420
QY	1460	ATAGTGATCAAAACGGATGTAAATTTATCAATTTTACACAAATTTGCTGAGCCGAGTGGAT	1519
DB	421	IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValAlaAspArgValAsp	440
QY	1520	GCAGAAGATGGACAGTATGATGTTTATGTTTATCGAAGACAGATGTGGAGCCGTTCTTAAA	1579
DB	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
QY	1580	GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAATG	1639
DB	461	ValValSerIleProLysGluThrTrpTyrAspLeuGluGluValLeuLeuGluGluMet	480
QY	1640	ACAGTTTTTGGGAACCGACTGTATTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACA	1699
DB	481	ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
QY	1700	CTATATATGTTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC	1759
DB	501	LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr	520
QY	1760	GGAAAGCGTGTGCTAGTGTGTGCTGCTGCCGAGACCTTACTGCTGCTGGATGTTCT	1819
DB	521	GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaLysPaspGlySer	540
QY	1820	GCATGTTCTCGCTATTTTCCACTGCTGCAAGAGCGCAGACAGAGATATAGAAT	1879
DB	541	AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn	560
QY	1880	GGAGACCACTGACTCTCTTCCAGACTTACACCATGATATACCATGCGCCACAGCCCT	1939
DB	561	GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro	580
QY	1940	GAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATTTTGGGAATCAGTCCGAAG	1999
DB	581	GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys	600

QY	2000	TCGCAGAGCGGCTGCTCTATTGGCAATTTCCAGAGCGGCAATAGAGAGCGAAAGAGAGAG	2059
Db	601	SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu	620
QY	2060	ATCAGAGTGGATGATCATATCATCATCAGCACAGATCAAGAGCCCTTCTGCTACGTAGTCTACAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln	640
QY	2120	CAGAAGATTTCAGGCAATTAACCTCTGCCATGCGGTGGAACATGGGTTTCATACAAACTCTT	2179
Db	641	GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu	660
QY	2180	CTTAAGGTAAACCTGGAGTCATTGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT	2239
Db	661	LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp	680
QY	2240	GATCATGGAGATGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	681	AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys	700
QY	2300	GTCTGTGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGATGGATGAG	2359
Db	701	ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
QY	2360	TTCTGTGAACAAGTTTGGAAAGGACCGGAAACAACGTGCGCAAGCCGAGACATACC	2419
Db	721	PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr	740
QY	2420	CCAGGGACAGTAAACAATGGAGCACTTACAGAAAAATAGAAAGTAGAACAACAGGAGG	2479
Db	741	ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg	760
QY	2480	ACCACCAATTTGAGAGCGCACCCAGGAGTGTC	2512
Db	761	ThrHisGluPheGluArgAlaProArgSerVal	771
RESULT 3			
ABG96413			
ID	ABG96413 standard; Protein; 771 AA.		
XX	AC ABG96413;		
XX	11-DEC-2002 (first entry)		
XX	Human ovarian cancer marker M473.		
XX	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
KW	central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
XX	histological type; carcinogenic; ovarian cancer marker.		
XX	Homo sapiens.		
XX	WO200271928-A2.		
XX	19-SEP-2002.		
XX	14-MAR-2002; 2002WO-US07826.		
XX	14-MAR-2001; 2001US-276025P.		
PR	14-MAR-2001; 2001US-276026P.		
PR	10-AUG-2001; 2001US-311732P.		
PR	19-SEP-2001; 2001US-323580P.		
PR	26-SEP-2001; 2001US-324967P.		
PR	26-SEP-2001; 2001US-325102P.		
PR	26-SEP-2001; 2001US-325149P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
XX	PA		
XX	XX		

QY 500 GATGAATGCAAGTGGCTGGAAAAGACATCTCTCAAGAATGTGCTAAATTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATCTGTAGCTGTGGAAACGGGGCTTTTTCATCAATTT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAATTCGACATCATCTCTGGAGCAATATTTTAAGCTGGAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGCGCGCTGGGAAGAGTCCATATGACCCCTAAAGCTGTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTAATCTCTGGAAGTCTGAGCTGATTTTATGGGGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTCTGGGCACCAACCAATCAGACAGACAGCAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTCAATTAAGTGGCCACCTCATCTCAGAGGTGCAAAATCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGACACTCTGGAAGACT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATATGACAGATGACTTTGGAGCGGCACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACAACTTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGC 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTGAAGCTGAGGATGATTCCTAATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPhelyAspProLys 320
QY 1160 AATCAGTTGTATATGGAGTGTTCAGACTTCCTCAGTAACTTTTCAAGGGATCAGCGGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGATATGATAGCATGAGTATGAGAGGGTGTTCCTTTGGTCCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCACGACAAACATTTGGTGGTTTTCAGTCTCAAGAGGACCTTCCTGATGATGCTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCATGTACATCCAGTGTTCCTTATCAAACTCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGATGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTTCTAAGGAGACTTGGTATGATTTTGAAGAGGTTCTGCTGGAAGAAATG 1639

Db 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCCGGAAACCGACTCTGCTATTTCACCAATGAGAGCTTTCACCAAGCAGCAAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATGCTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGTGGATGGTCTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCACAGACCAAGATATTAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCACTGTTTCAGACTTACACCATGATATCACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspPheHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGAATCATCTATGCTGTAGAGAATAGTAGACATTTTTCGAAATGCGAGTCCGAAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGTGCTTATTTGGCAATTCAGAGGCGAAATGAAGACGCAAGAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGSCCTTCTGCTACGTAGTCTCAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTAACTCCCTGGGAAGTCATTGACACAGAGCATTTGGNAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATCATCGAGATGCTCTTAAGACCAAGAAATGTCCTAATAGCATGACATGACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGCTCAGCTCATCAACCCCAATCTCAACACGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAAGGACCGAAACCACTCGGCAAGGCCGAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAAACAGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGGCACCCAGGAGTGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
AAY21264
ID AAY21264 standard; Protein; 796 AA.
XX AC AAY21264;
XX AC AAY21264;
XX DT 22-JUL-1999 (first entry)
XX DE Human semaphorin III wild type protein fragment 1.
XX

KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Homo sapiens.

XX

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

XX corresponding protein mutations - used to diagnose cancer and

XX neurological diseases, particularly Alzheimer's disease, and also

XX for treatment and prevention with specific ribozymes or wild-type

XX RNA

XX Disclosure; Figure 16; 258pp; English.

XX

XX This invention describes a novel method for the diagnosis of a disease

XX caused by, or associated with, an RNA molecule that has a frameshift

XX mutation. The method is used to diagnose age-related diseases, especially

XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's

XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,

XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II

XX and many others listed) or susceptibility to these disorders. The method

XX allows a definitive diagnosis of Alzheimer's disease in living patients,

XX at an early stage. It is based on the observation that disease may be

XX caused by mutations in RNA rather than DNA. The invention describes the

XX use of neuronal system RNA molecules, specifically proteins including

XX beta-amyloid precursor protein (beta-APP), the microtubule associated

XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule

XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,

XX neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic

XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma

XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group

XX protein-C (HMGP-C) and neuroendocrine specific protein A.

XX

XX SQ Sequence 796 AA;

Alignment Scores:

Pred. No.:	Score:	Length:
0	437.00	796
Percent Similarity:	100.00%	Matches: 437
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	48.83%	Mismatches: 0
DB:	19	Indels: 0
		Gaps: 0

US-09-774-490-1 (1-2709) x AAY21264 (1-796)

QY	194	TGCAGCATGGCTGGTAACTAGGATTCGTCTCTTTCTGGGGAGTATTACTACAGCA	253

DB	4	CysSerMetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAla	23

QY	254	AGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGCGTGAATATTCCTACAAAGAA	313
DB	24	ArgAlaAsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGlu	43
QY	314	ATGTTGGAAATCCAAATGTGATCCTTCAATGGCTTGGCCACAGCTCCAGTTATCAT	373
DB	44	MetLeuGluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHis	63
QY	374	ACCTTCTCTTTGGATGAGAAACGAGTACGCTGTATGTTGGAGCAAAAGATCACATATTT	433
DB	64	ThrPheLeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePhe	83
QY	434	TCATTGACCTGGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTACACC	493
DB	84	SerPheAspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThr	103
QY	494	AGAGAGATGAATCAAGTGGCTGGAAAGACACATCTCTGAAGATGTGCTAAATTTTATC	553
DB	104	ArgArgAspGluCysLysIleValGlyLysAspIleLeuLysGluCysAlaAsnPheIle	123
QY	554	AAGTACTTAAAGCATATATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTTCAT	613
DB	124	LysValLeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHis	143
QY	614	CCAATTTGCACCTACATTTGACATCATCTCTGAGGACAATATTTTAAAGCTGGAG	673
DB	144	ProIleCysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGlu	163
QY	674	AACTCACATTTTGAACGGCTGGGAGAGTCCATATGACCTTAAGCTCTGACACA	733
DB	164	AsnSerHisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAla	183
QY	734	TCCCTTTTAAATAGATGAGAAATTAATCTCTGAACTGCAGCTGATTTTATGGGGAGAC	793
DB	184	SerLeuLeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAsp	203
QY	794	TTTGCTATCTTCCGAACCTTTGGGCACACACCAATCAGGACAGACAGCATGATTC	853
DB	204	PheAlaIlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSer	223
QY	854	AGTGGCTCAATGATCCAAAGTTTCAATGATGCCACCTCATCTCAGAGAGTGCATCTCT	913
DB	224	ArgTrpLeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnPro	243
QY	914	GAACATCAAAAGTATATCTTTCTCGTCAAAATGCAATAGATGGAGAACACTCTCGA	973
DB	244	GluAspAspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGly	263
QY	974	AAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGATTTTGGAGGGCCACAGAGT	1033
DB	264	LysAlaThrHisAlaAArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSer	283
QY	1034	CTGCTGATTAATGGACAACATCTCTCAAGCTCGTCTGATTTCTCAGTCCAGGTCCA	1093
DB	284	LeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyPro	303
QY	1094	AATGGCATTCACACTCATTTTGTGAATGAACTGAGATGTATTCTTAATGAACCTTTAAAGAT	1153
DB	304	AsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPhenylAsp	323
QY	1154	CCTAAAAATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAAACATTTTCAAGGGATCA	1213
DB	324	ProLysAsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySer	343
QY	1214	GCCGTGTATGTATACCATGATGATGAGAGGGGTGTTCTTGGTCCATATGCCAC	1273
DB	344	AlaValCysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHis	363
QY	1274	AGGATGACCCCACTATCAATGGTGCCTTTATCAAGAGAGTCCCTTATCCACGGCA	1333
DB	364	ArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgPro	383
QY	1334	GGAACTTGTCCACGCAAAACATTTGGTGGTGTGACTCTCAAGAGGACCTTCTCTGATGAT	1393

Db 384 GlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAsp 403
 Qy 1394 GTTATAACCTTTGCAAGAAGTCAATCCAGCATGTACATCCAGTGTTCCTATGACAAT 1453
 Db 404 ValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsn 423
 Qy 1454 CGCCCAATAGTCATCAAAACGATGTAAATTAATCAATTTACACAAATGTC 1504
 Db 424 ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleVal 440

RESULT 5

AAR74175

ID AAR74175 standard; Protein; 477 AA.

XX

AC AAR74175;

XX

DT 01-NOV-1995 (first entry)

XX

DE Human collapsin.

XX

KW Collapsin; antibody; therapy.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Binding-site 9..19

FT Binding-site /note= "antibody binding site"

FT Binding-site 51..65

FT Binding-site /note= "antibody binding site"

XX

XX US5416197-A.

PN

XX 16-MAY-1995.

PD

XX 15-OCT-1993; 93US-0136922.

XX

XX 15-OCT-1993; 93US-0136922.

XX

XX (UYPE-) UNIV PENNSYLVANIA.

XX

XX Luo Y, Raper JA;

XX

XX WPI; 1995-193478/25.

XX

XX N-PSDB; A9Q92331.

XX

PT New antibody to human collapsin - used to inhibit the activity of
 PT collapsin, to induce neurite out-growth and to treat individuals with
 PT nerve damage.

XX

XX Claim 2; Columns 15-18; 11pp; English.

XX

CC An antibody capable of specifically binding at least a portion of
 CC the collapsin protein can be used to purify human collapsin and
 CC to inhibit the activity of the protein. It can be used to induce
 CC neurite outgrowth by neuronal cells and to treat individuals
 CC suffering from nerve damage.

XX

SQ Sequence 477 AA;

XX

Alignment Scores:

Pred. No.: 0 Length: 477
 Score: 434.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.49% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR74175 (1-477)

Qy 905 GACATCTCGAGATGACAAATGATACCTTTCTTCCTGAAATGCAATGATGGAGAA 964

Db 4 AspAsnProGluAspAspLysValTyrPhePheArgGluAsnAlaIleAspGlyGlu 23

Qy 965 CACTCTGAAAGCTACTCACGCTAGATAGTGCAGATATCAAGAATGACTTTTGGAGG 1024
 Db 24 HisSerGlyLysAlaThrHisAlaArgLleGlyGlnIleCysLysAsnAspPheGlyGly 43
 Qy 1025 CACAGAAGTCTGGTGAATAAATGACACCAATTCCTCAAGCTCGTCTGATTGTCACGTG 1084
 Db 44 HisArgSerLeuValAsnLysTyrThrPheLeuLysAlaArgLeuLleCysSerVal 63
 Qy 1085 CCAGGTCAAATGGCATTGACACTCATTTTTCATGAACTGCAGGATGATTTCTTAATGAAC 1144
 Db 64 ProGlyProAsnGlyLleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 83
 Qy 1145 TTTAAAGATCCTAAAAATCCAGTTGTATATCGAGTGTTCAGACTTCACAGTAACATTTTC 1204
 Db 84 PheLysAspProLysAsnProValValTyrGlyValPheThrThrSerSerAsnLlePhe 103
 Qy 1205 AAGGATCAGCCGTGTGTATGTATAGCATGATGTGAGAAGGGTGTTCCTTGGTCCA 1264
 Db 104 LysGlySerAlaValCysMetTyrSerMetSerAspValArgValPheLeuGlyPro 123
 Qy 1265 TATCCCAACAGGATGACACCACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTAT 1324
 Db 124 TyrAlaHisArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyr 143
 Qy 1325 CCACGGCCAGGAACCTTGTCCAGCAAAACATTTTGGTGTGTTCAGTCTTACAAAGACCTT 1384
 Db 144 ProArgProGlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeu 163
 Qy 1385 CCTGATGATGTTATACCTTTGCAAGAAAGTCATCCAGCCATGTACAATCCAGTGTTCCT 1444
 Db 164 ProAspAspValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPhePro 183
 Qy 1445 ATGAACAATGCCCAATAGTATCAAAACGGATGTAATTAATTAATTAACAATAATGTC 1504
 Db 184 MetAsnAsnArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleVal 203
 Qy 1505 GTAGACCGAGTGGATGACAGAAGATGACAGTATGATGTATGTTTATCGGAACAGATGTT 1564
 Db 204 ValAspArgValAspAlaGluAspGlyGlnTyrAspValMetPheLleGlyThrAspVal 223
 Qy 1565 GGGACCGTCTTAAAGTAGTTTCAATTCCTAAGAGAGACTTGGTATGATTTAGAGAGGTT 1624
 Db 224 GlyThrValLeuLysValValSerIleProLysGluThrTrpTyrAspLeuGluVal 243
 Qy 1625 CTGCTGAAGAATGACAGTTCGGAACCGACTGCTATTTTCAGCAATGGAGCTTCC 1684
 Db 244 LeuLeuGluGluMetThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSer 263
 Qy 1685 ACTAAGCAGCAACAACCTATATATTGTTCAACGCTGGGTTGCCAGCTCCCTTTACAC 1744
 Db 264 ThrLysGlnGlnLeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHis 283
 Qy 1745 CGGTGTGATATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGT 1804
 Db 284 ArgCysAspIleTyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 303
 Qy 1805 GCTTGGATGTTCTGCATGTTCTCGCTATTTTCCACTGCAAGAGACCGCAGACGCA 1864
 Db 304 AlaTrpAspGlySerAlaCysSerArgTyrPheProThrAlaLysArgThrArgArg 323
 Qy 1865 CAAGATATAAGAAATGGAGACCCACTGACTCCTGTTTCAGACTTACACCATGATAATCAC 1924
 Db 324 GlnAspIleArgAsnGlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHis 343
 Qy 1925 CATGGCCACACCCCTGAAGAGAGAATCATCTATGGTGTAGAGATAGTAGACATTTTGT 1984
 Db 344 HisGlyHisSerProGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeu 363
 Qy 1985 GAATGAGTCCGAGTCCGAGAGACGCTGCTGTTATGGCAATTCAGAGCGGAAATCAA 2044
 Db 364 GluCysSerProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGlu 383

QY 2045 GAGCGAAAGACAGATCAGAGTGGATGATCATATCATCATCAGACAGATCAAGGCTTCTG 2104
 Db GluArgLysGluGluIleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeu 403
 QY 2105 CTACCTAGTCTACACAGAGATTGAGGCAATTAACCTCTGCATGCGGTGGAACATCGG 2164
 Db LeuArgSerLeuGlnGlnLysAspSerGlyAsnTyrLeuCyHisAlaValGluHisGly 423
 QY 2165 TTCATACAACTCTTCTTAAGTAAACCTCGGAAGTCATTGAC 2206
 Db PheIleGlnThrLeuLeuLysValThrLeuGluValIleAsp 437
 RESULT 6
 ID ABG54800 standard; Peptide; 151 AA.
 XX
 AC ABG54800;
 XX
 DT 25-FEB-2003 (first entry)
 DE Human liver peptide, SEQ ID No 33448.
 DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 PS Claim 27; SEQ ID No 33448; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 151 AA;
 Alignment Scores:
 Pred. No.: 7,288-145 Length: 151
 Score: 151.00 Matches: 151
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 16.87% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x ABG54800 (1-151)
 QY 2060 ATCAGAGTGGATGATCATATCATCATCAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
 Db 1 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 20
 QY 2120 CAGAGGATTCAGACCAATTACCTCTGCCATCGGTGGAACATGGTTCATACAACTCTT 2179
 Db 21 GlnLysAspSerGlyAsnTyrLeuCyHisAlaValGluHisGlyPheIleGlnThrLeu 40
 QY 2180 CTTAAGGTAAACCTCGGAAGTCATTGACACAGAGCATTTGGGAAGAACTTCTTCATAAAGAT 2239
 Db 41 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 60
 QY 2240 GATGATGGAGATGCTCTTAAGACCAAGAAATGTCCAATACATGACACCTAGCCAGAAG 2299
 Db 61 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 80
 QY 2300 GTCTGTACAGAGACTTCATGACGATCATCAACACCCCATCTCAACACCATGATCAG 2359
 Db 81 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 100
 QY 2360 TTCTGTGAACAAAGTTTGGAAAAGGACCGAAAACAAAGCTCGGCAAGCCAGGACATACC 2419
 Db 101 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 120
 QY 2420 CCAGGGAACAGTAACAAATGGAAGCATTCAAGAAATAAGAAAGGTAGAAAACAGGAGG 2479
 Db 121 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 140
 QY 2480 ACCACCAAGTTTGAGGGGACCCAGGAGTGTC 2512
 Db 141 ThrHisGluPheGluArgAlaProArgSerVal 151
 RESULT 7
 ID ABG42930
 XX
 AC ABG42930;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 32595.
 DE Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Rudik syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPT; 2002-114183/15.
XX DR
XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX Claim 27; SEQ ID No 32595; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic,
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a peptide/protein
XX encoded by a single exon probe of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 151 AA;

Alignment Scores:
Pred. No.: 7.28e-145 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x ABG42930 (1-151)

QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCCCTTCGTACGTAGTACAA 2119
Db " 1 IleArgValAspHisIleAlaThrAspGlnGlyLeuLeuLeuAArgSerLeuGln 20

QY 2120 CAGNAGGATTCAGCAATTACCTCTCGCATCGGTGGAACTGGGTTCAATCAAACTCTT 2179
Db 21 GlnLysAspSerGlyAsnTrpLeuCyHisAlaValGluHisGlyPheIleGlnThrLeu 40

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ Sequence 81 AA;

XX Alignment Scores:

Pred. No.:	2,81e-73	Length:	81
Score:	81.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.05%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21293 (1-81)

QY 2286 CACCTAGCCAGAGGCTGTGTACAGAGCTTCATGACCTCATCAACCCCAATCTCA 2345
 Db 1 HistidAlaArgArgSerCylThreIunThrSerCysSerSerThirProIleSer 20
 QY 2346 ACACGATGATGATGTTCTGTGAACAAGTTGAAAAAGGAGCCGAAACAACTGGCAAA 2405
 Db 21 ThrArgTpmSerSerSerValAsnLysPheGlyLysGlyThGluAsnAsnValGlyLys 40
 QY 2406 GGGCGAGCATATCCCGAGGAGCACTAACAATGAGACGACTTACAGAAATAAGAAAG 2465
 Db 41 GlyIAsnIlePleProGlnGlyThrValThraGngLysSerThrTylLysLysIleArgLys 60
 QY 2466 GTAGAAAGAGAGAGAGCCAGCAATTTGAGAGGGAGCCAGAGATGCTGAGCTGATAC 2525
 Db 61 ValGluThrGlyGlyProThrAsnLeuArgGlyHisProGlyValSerGluLeuHisTyr 80

QY 2526 CTC 2528

Db 81 Leu 81

RESULT 9

AAY21288 standard; Protein; 77 AA.

AC AAY21288;

DT 22-JUL-1999 (first entry)

XX Human semaphorin III mutant protein fragment 21.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;

XX frameshift mutation; age-related disease; neurodegenerative disorder;

XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

XX Huntington's disease; multiple sclerosis; alcoholic liver disease;

XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;

XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;

XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;

XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;

XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSF-A;

XX high mobility group protein-C; neuroendocrine specific protein A.

OS Homo sapiens.
 XX
 XX WO9845322-A2.
 XX
 XX 15-OCT-1998.
 XX
 XX 02-APR-1998; 98WO-IB00705.
 XX
 XX 10-APR-1997; 97US-0043163.
 XX
 XX (UUYT-) RIJCKSINIV UTRECHT.
 XX (ROYA-) ROYAL NETHERLANDS ACADEMIES OF SCI.
 XX (UUYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 XX Burdach JPH, Grosfeld FG, Van Leeuwen FW;
 XX
 XX WPT: 1998-609901/51.
 XX
 XX N-PSDB: AAX75767.
 XX
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 XX corresponding protein mutations - used to diagnose cancer and
 XX neurological diseases, particularly Alzheimer's disease, and also
 XX for treatment and prevention with specific ribozymes or wild-type
 XX RNA

Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ Sequence 77 AA;

XX Alignment Scores:

Pred. No.:	3.47e-69	Length:	77
Score:	77.00	Matches:	77
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.60%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21288 (1-77)

QY 1641 CAGTTTTCGGGAAACCGACTCTATTTCAGCAATGAGCTTTCACCTAAGAGCAACAAC 1700
 Db 1 GlnPhePheGlyAsnArgLeuLeuPheGlnGlnIlnTrpSerPheProLeuSerSerAsnAsn 20
 QY 1701 TATATATTGTTTCAACGGCTGGGTTGGCCAGCTCCCTTACACGGGTGTATATTACG 1760
 Db 21 TyrIleLeuValGlnArgLeuGlyLeuProSerSerLeuTyrThrGlyValIlePheThr 40
 QY 1761 GGAAGCGTGTGAGTGTGAGTGTGCTGGCCGAGACCCCTTACTGTGCTGGAGATGTTCTG 1820
 Db 41 GlyLysArgValIleLeuSerValAlaSerProGluThrLeuThrValLeuGlyMetValLeu 60
 QY 1821 CATGTTCTCGCTATTTCCTCACTGCACAGAGCCACACAGACGACAGATA 1871
 Db 61 HisValIleAlaIlePheProLeuGlnArgSerPalaGlnAspAlaPylsIle 77

RESULT 10
AAV21291
ID AAV21291 standard; Protein; 74 AA.
XX
AC AAV21291;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human semaphorin III mutant protein fragment 24.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98MO-IB00705.
XX
PR 10-APR-1997; 97US-0043163.
XX
PA (UYUT-) RIJSDUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
PI Burdach JPH, Grosveld FG, Van Leeuwen FW;
XX
DR MPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
PS Disclosure; Figure 16; 258bp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-APP), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
SQ Sequence 74 AA;
XX
Alignment Scores: 4.07e-66 Length: 74
Pred. No.:

Score: 74.00 Matches: 74
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.27% Indels: 0
DB: 19 Gaps: 0
US-09-774-490-1 (1-2709) x AAV21291 (1-74)
QY 1965 AGAATAGTAGCATTCTTTTGGATGACGACGCGAGTGGCGAGCGCTGCTATTGGC 2024
DB 1 ArgilleValAlahisPhetripasnlavalAgsersrArgAluArgtrpserillegly 20
QY 2025 AATTCGAGCGCGGAATGAGAGCGAAGAAAGAGATCGAGTGCATGATCATCATCA 2084
DB 21 AnserArngglyGlmetllyserlulysrAsgserGluTrpMetleleeser 40
QY 2085 GCACGATCAGAGCCCTTGCTGCTAGCTAGTCTTCAACAGAGAGATTGCGCAATTACCTCT 2144
DB 41 GlylnllylValAlphecYerValValTyraAsnArgArgllleGlnAlaIleThrser 60
QY 2145 GCCATGCGGTGGACATGGGTTCATACAAACCTCTTAAGG 2186
DB 61 AlameArgrTpAsmcellyserTyrysleuPheleuArg 74
RESULT 11
AAV21274
ID AAV21274 standard; Protein; 69 AA.
XX
AC AAV21274;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human semaphorin III mutant protein fragment 7.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98MO-IB00705.
XX
PR 10-APR-1997; 97US-0043163.
XX
PA (UYUT-) RIJSDUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
PI Burdach JPH, Grosveld FG, Van Leeuwen FW;
XX
DR MPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
PS Disclosure; Figure 16; 258bp; English.
XX

CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HRP-1, high mobility group
CC protein-C (HMGp-C) and neuroendocrine specific protein A.

SQ Sequence 69 AA;

Alignment Scores:

Pred. No.:	5,3e-61	Length:	69
Score:	69.00	Matches:	69
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.71%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21274 (1-69)

QY 747 ATGAGAAATTAATCTCTGGAAGTGAAGTATTTATGGGAGAGACTTGTATCTTCC 806
DB 1 MetGluAunfYrThreLgluLnuLnuLleuLtrpGluThreLnuLnuSerSer 20
QY 807 GAACCTTTGGGACACACCAATCAGACAGACAGATGATTCAGTGGCTCAATG 866
DB 21 GluLeuLnuGlyThrThrThrGlnSerGlyGlnSerSerMetLeuProGlyGlySerMet 40
QY 867 ATCCAAAGTTCATTTAGTGGCCACCTCATCTCAGACAGATCAATCTGAAGTGAAGAAG 926
DB 41 IleGlnSerSerLeuValProThrSerSerGlnArgValThrIleLnuLnuMetThrIys 60
QY 927 TATACCTTTTCTCCGTGAATAATGCAA 953
DB 61 TyrThrPheSerSerValIysMetGln 69

RESULT 12

AAY21272 standard; Protein; 64 AA.

ID AAY21272;

AAV21272;

22-UTU-1999 (first entry)

Human semaphorin III mutant protein fragment 5.

Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
frameshift mutation; age-related disease; neurodegenerative disorder;
Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
Huntington's disease; multiple sclerosis; alcoholic liver disease;
diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
neurofilament-F; presentin I; presentin II; cellular tumour antigen;
glial fibrillary acidic protein; GFAP; p53; semaphorin III; HRP-1;
bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
high mobility group protein-C; neuroendocrine specific protein A.

Synthetic.

Homo sapiens.

WO9845322-A2.

PD 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJCKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

XX corresponding protein mutations - used to diagnose cancer and

XX neurological diseases, particularly Alzheimer's disease, and also

XX for treatment and prevention with specific ribozymes or wild-type

XX RNA

XX Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HRP-1, high mobility group
CC protein-C (HMGp-C) and neuroendocrine specific protein A.

SQ Sequence 64 AA;

Alignment Scores:

Pred. No.:	6,91e-56	Length:	64
Score:	64.00	Matches:	64
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.15%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21272 (1-64)

QY 534 AAGATGTGTAATTTATCAAGTAAAGCATATATACAGACTTCATTTGACGCT 593

DB 1 LysAsnValIleuLleSeSerArgIleuArgHisIleLnuLnuThrCysThrPro 20

QY 594 GTGAAAGCGGCGCTTTTCATTCATTTGACCTACATTTGAATTTGACATCATCTGAGG 653

DB 21 ValGluArgGlyLeuPheIleGlnPheAlaProThrLeuLysLeuAspIleIleLnuArg 40

QY 654 ACAATATTTTAAAGTGAAGATTCACATTTTGAAGACCGCGTGGGAAGTCCATATG 713

DB 41 ThrIlePheLysSerThrPArgThrHisIleLnuLnuThrAlaValGlyArgValHisMet 60

QY 714 ACCCTAAGCTGC 725

DB 61 ThrLeuSerCys 64

RESULT 13

AAY21299 standard; Protein; 57 AA.

XX AAY21299;
 AC 22-JUL-1999 (first entry)
 XX Human semaphorin III mutant protein fragment 32.
 DE Human semaphorin III mutant protein fragment 32.
 XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX Synthetic.
 OS Homo sapiens.
 XX MO9845322-A2.
 PN 15-OCT-1998.
 PD 02-APR-1998; 98WO-IB00705.
 PF 10-APR-1997; 97US-0043163.
 PR (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX Butbach JPH, Grosveld FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX Disclosure; Figure 16; 258pp; English.
 PS This invention describes a novel method for the diagnosis of a disease
 XX caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC protein Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX SQ Sequence 57 AA;
 Alignment Scores:
 Pred. No.: 1e-48 Length: 57
 Score: 57.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.37% Indels: 0

DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21299 (1-57)
 QY 217 GATGTCTGCTTTTCTGGGAGTATTACTTACAGCAAGCAAACTACAGATGGAA 276
 DB 1 AspCysIeuSerPheIeuNylYserIleThrYIserIlySerIlySerIlySerIlyIutpGlu 20
 QY 277 GAACATATGCCAAGGCTGAATTTATCTTACAAAGAAATGTTGAATCCAAATGTGAT 336
 DB 21 GluGlnCysAlaIysAlaIleIleIleuGlnArgAsnValGlyIleGlnIlyCysAsp 40
 QY 337 CACTTTCATGGCTTGGCCACAGCTCCAGTTATCATCTTCTTTTGA 387
 DB 41 HispneGlnItrpIeuNylYlGlnIlnuGlnIeuSerIlyLeuProPheIly 57
 RESULT 14
 AAY21271
 ID AAY21271 standard; Protein; 85 AA.
 XX AAY21271;
 AC 22-JUL-1999 (first entry)
 XX Human semaphorin III mutant protein fragment 4.
 DE Human semaphorin III mutant protein fragment 4.
 XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX Synthetic.
 OS Homo sapiens.
 XX MO9845322-A2.
 PN 15-OCT-1998.
 PD 02-APR-1998; 98WO-IB00705.
 PF 10-APR-1997; 97US-0043163.
 PR (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX Butbach JPH, Grosveld FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX Disclosure; Figure 16; 258pp; English.
 PS This invention describes a novel method for the diagnosis of a disease
 XX caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,

CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including the
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC protein Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin 1, presenilin 2, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGp-C) and neuroendocrine specific protein A.

XX Sequence 85 AA;

Alignment Scores:

Pred. No.:	1,05e-46	Length:	85
Score:	55.00	Matches:	55
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.1%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21271 (1-85)

QY 366 GTTATCATACCTTCTTTGGATGAGGAACGAGTACGCTGATGTTGAGCAAGATC 425

DB 31 ValIleIleProSerPheTrpMetArgAangIValGIyCyMetLeuGIuGIaArgIle 50

QY 426 ACATATTTTCATCGACCGTGAATATATCAAGATTTTCAAAAGATGTGCGCCATAT 485

DB 51 ThrTrpPheHisSerThrTrpLeuIleSerArgIlePheIleValArgLeuGlyGlnIlyr 70

QY 486 CTTACACCAAGAGATGATGATCAAGTGGCTGCAAAAGACATCC 530

DB 71 LeuThrProGIuGIuMetAsnIaSerGIyLeuGIuIlyThrSer 85

RESULT 15

ABG54764 ID ABG54764 standard; Peptide; 57 AA.

XX AC ABG54764;

DT 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 33412.

DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

PN WO200157273-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human adult liver -

XX

PS Claim 27; SEQ ID No 33412; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SRNP) (i) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (i) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG5930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 57 AA;

Alignment Scores:

Pred. No.:	1.23e-44	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x ABG54764 (1-57)

QY 311 GAATGTTGGAATCCCAAGATGATCACTTCAATGCTTGCCACAGCTCCAGTTAT 370

DB 5 GluMetLeuGIuSerAsnAsnValIleThrPheAsnGIyLeuAlaAsnSerSerIlyr 24

QY 371 CATACCTTCCTTTTGGATGAGGACGAGTACGCTGATGTTGGCAAGATCACTA 430

DB 25 HisThrPheLeuLeuAspGIuGIuArgSerArgLeuIlyValGIyAlaIlyAspHisIle 44

QY 431 TTTTCATCGACCGTGAATATATCAAGATTTTCAAAAG 469

DB 45 PheSerPheAspLeuValIleIlyAspPheGIuIlyS 57

RESULT 16

ABR39696 ID ABR39696 standard; Peptide; 57 AA.

XX AC ABR39696;

DT 04-FEB-2002 (first entry)

DE Peptide #7202 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

PN WO200157277-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 32331; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WFO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 57 AA;

Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x ABB39696 (1-57)

OY 311 GAATGTTGAATCCAAAGATGATCATTTCATGCTTGCCCAACAGCTCCAGTTAT 370
DB 5 GllmetleuGluserAenValillethPheAenGlyLeuAlaAenSerSeryr 24
OY 371 CATACCTTCCTTTTGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HlsthPheLeuAenSpGluGluArgSerArgLeuYrValGlyAlaLysAspHisle 44
OY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAenIlelyAspPheGlnlys 57

RESULT 17
AAM60414
ID AAM60414 standard; Protein; 57 AA.
XX
AC AAM60414;

DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32519.

XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.

XX
PN WO200157275-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.

XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 32519; 650pp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 57 AA;

Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAM60414 (1-57)

OY 311 GAATGTTGAATCCAAAGATGATCATTTCATGCTTGCCCAACAGCTCCAGTTAT 370
DB 5 GllmetleuGluserAenValillethPheAenGlyLeuAlaAenSerSeryr 24
OY 371 CATACCTTCCTTTTGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HlsthPheLeuAenSpGluGluArgSerArgLeuYrValGlyAlaLysAspHisle 44
OY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAenIlelyAspPheGlnlys 57

RESULT 18
AAM73050
ID AAM73050 standard; Protein; 57 AA.
XX
AC AAM73050;

DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33356.

XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.

XX
PN WO200157276-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.

XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT
XX
PS Example 4; SEQ ID NO: 33356; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x AAM73050 (1-57)
QY 311 GAAATGGTGAATCCAAATGATGATCACTTCAATGGCTTGCCAAAGCTCCAGTTAT 370
DB 5 GtmetleugluserashsnvalletthrPheanglyleuAlasnserserTyr 24
QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTGAGCAAGATCACATA 430
DB 25 HstthrPheleuleuAspGluGluArgSerArgleuTyrValGlyAlaIysAspHisIle 44
QY 431 TTTTCATTGCACCTGCTTATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleLysAspPheGlnLys 57
XX
RESULT 19
AAM33272
ID AAM33272 standard; Protein; 57 AA.
XX
AC AAM33272;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #7309 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 33541; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see A11315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x AAM33272 (1-57)
QY 311 GAAATGGTGAATCCAAATGATGATCACTTCAATGGCTTGCCAAAGCTCCAGTTAT 370
DB 5 GtmetleugluserashsnvalletthrPheanglyleuAlasnserserTyr 24
QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTGAGCAAGATCACATA 430
DB 25 HstthrPheleuleuAspGluGluArgSerArgleuTyrValGlyAlaIysAspHisIle 44
QY 431 TTTTCATTGCACCTGCTTATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleLysAspPheGlnLys 57
XX
RESULT 20
ABG42894
ID ABG42894 standard; Peptide; 57 AA.
XX
AC ABG42894;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 32559.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX

Db 21 SerHisCysLysGluThrHisIleThrArgTyrLysLeuTyrPargProThrAspSer 40
 QY 1897 CTGTCAGACTTACCA 1914
 |||||
 Db 41 LeuPheArgLeuThrPro 46
 RESULT 25
 AAY21322
 ID AAY21322 standard; Protein; 42 AA.
 XX
 AC AAY21322;
 XX
 XX 22-JUL-1999 (first entry)
 DE Human semaphorin III mutant protein fragment 55.
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSF-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 XX Homo sapiens.
 XX
 XX WO9845322-A2.
 XX
 XX 15-OCT-1998.
 XX
 XX 02-APR-1998; 98WO-IB00705.
 XX
 XX 10-APR-1997; 97US-0043163.
 XX
 XX (UYUT-) RIJKSUNIV UTRECHT.
 XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 XX (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
 PI
 DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 XX Discloure; Figure 16, 258pp; English.
 XX
 XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ: Sequence 42 AA;
 Alignment Scores:
 Pred. No.: 2,24e-33 Length: 42
 Score: 42.00 Matches: 42
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.69% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21322 (1-42)
 QY 1240 TGTGAGAGGGTGTCTTGGTCATATGCCACAGAGTGAGCCCACTATCATGGGT 1299
 |||||
 Db 1 CysGluYrGlyValProTyrPserIleCysProGlnGlyTyrThrGlnLeuSerMetCly 20
 QY 1300 GCCTTATCAAGAGAGTCCCTATCCAGCGCCAGAACTGTCCCAAAACATTGG 1359
 |||||
 Db 21 AlaLeuSerArgLysSerProLeuSerThrAlaArgAsnLeuSerGlnGlnIleTyr 40
 QY 1360 TGGTTT 1365
 |||||
 Db 41 TrpPhe 42
 RESULT 26
 ABG55921
 ID ABG55921 standard; Peptide; 41 AA.
 XX
 XX ABG55921;
 XX
 XX 25-FEB-2003 (first entry)
 XX
 XX Human liver peptide, SEQ ID No 34569.
 XX
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 XX Homo sapiens.
 XX
 XX WO200157273-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00664.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 DR WPI; 2001-488898/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 XX Claim 27; SEQ ID No 34569; 658pp; English.
 XX
 XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG59930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from Wfipo at ftp.wfipo.int/pub/published_pct_sequences.

XX
SQ Sequence 41 AA;

Alignment Scores:

Pred. No.:	2,37e-32	Length:	41
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.58%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x ABG55921 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGATGAGAGATAGACATTTTGGATGAGTCGG 1996

DB 1 PROGLUGLARGILELIERGLYValGluAnsSerThrPheLeuGluCysSerPro 20

QY 1997 AAGTCGACAGAGCGCTGCTTATTGGCAATTCCAGAGCGGAATGAAGCGAAAGAA 2056

DB 21 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAnGluGluArgTyrGlu 40

QY 2057 GAG 2059

DB 41 Gln 41

RESULT 27

ID AAM61385
ID AAM61385 standard; Protein; 41 AA.

XX AAM61385;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33490.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 33490; 650bp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.

XX
SQ Sequence 41 AA;

Alignment Scores:

Pred. No.:	2,37e-32	Length:	41
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.58%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAM61385 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGATGAGAGATAGACATTTTGGATGAGTCGG 1996

DB 1 PROGLUGLARGILELIERGLYValGluAnsSerThrPheLeuGluCysSerPro 20

QY 1997 AAGTCGACAGAGCGCTGCTTATTGGCAATTCCAGAGCGGAATGAAGCGAAAGAA 2056

DB 21 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAnGluGluArgTyrGlu 40

QY 2057 GAG 2059

DB 41 Gln 41

RESULT 28

ID ABG44063
ID ABG44063 standard; Peptide; 41 AA.

XX ABG44063;

DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 33728.

XX Human; single exon probe; asthma; lung cancer; COPD, ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

PD 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2002-114183/15.

Db 1 LysGluCysAlaAsnPhelIleValLeuLysAlaTyrAsnGlnThrHisLeuTyrAla 20
QY 593 TGTGAAACGGGGCTTTTCATCCAAATTGGACCTGACATT 631
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 34
AAM73166
ID AAM73166 standard; Protein; 33 AA.
AC AAM73166;
DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 33472.
DE Human bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma.
XX Homo sapiens.
OS
XX MO200157276-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00668.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 33472; 658bp + Sequence Listing; English.
XX PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX SQ Sequence 33 AA;
XX

Alignment Scores:
Pred. No.: 3,67e-24 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAM73166 (1-33)

QY 533 AAGAAATGTGCTAATTTCATCAAGGTAAGCATATATACAGCTGCTGTAAGCC 592
DB 1 LysGluCysAlaAsnPhelIleValLeuLysAlaTyrAsnGlnThrHisLeuTyrAla 20
QY 593 TGTGAAACGGGGCTTTTCATCCAAATTGGACCTGACATT 631
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 35
AAM33379
ID AAM33379 standard; Protein; 33 AA.
XX
XX AAM33379;
AC
XX
XX 17-OCT-2001 (first entry)
DT
XX
XX Peptide #7416 encoded by probe for measuring placental gene expression.
DE
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
OS
XX MO200157272-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00663.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID NO 33648; 654bp; English.
XX PS
XX The present invention relates to single exon nucleic acid probes (SENPs;
CC see A131315-A151546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
XX SQ Sequence 33 AA;
XX

Alignment Scores:
Pred. No.: 3,67e-24 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAM33379 (1-33)

QY 533 AAGAAATGTGCTAATTTCATCAAGGTAAGCATATATACAGCTGCTGTAAGCC 592
DB 1 LysGluCysAlaAsnPhelIleValLeuLysAlaTyrAsnGlnThrHisLeuTyrAla 20
QY 593 TGTGAAACGGGGCTTTTCATCCAAATTGGACCTGACATT 631
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 36
ABG43013
ID ABG43013 standard; Peptide; 33 AA.
AC ABG43013;
XX
XX

19-AUG-2002 (first entry)
Human peptide encoded by genome-derived single exon probe SEQ ID 32678.
Human; single exon probe; asthma; lung cancer; COPD; ILD;
chronic obstructive pulmonary disease; interstitial lung disease;
familial idiopathic pulmonary fibrosis; neurofibromatosis;
tuberosclerosis; Gaucher's disease; Niemann-Pick disease;
Hernandez-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
primary ciliary dyskinesia; pulmonary hypertension;
hyaline membrane disease.
Homo sapiens.
WO200186003-A2.
15-NOV-2001.
30-JAN-2001; 2001WO-US00665.
04-FEB-2000; 2000US-180312P.
26-MAY-2000; 2000US-207456P.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0633366.
21-SEP-2000; 2000US-234687P.
27-SEP-2000; 2000US-236359P.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2002-114183/15.
Spatially-addressable set of single exon nucleic acid probes, used to
measure gene expression in human lung samples -
Claim 27; SEQ ID NO 32678; 634bp; English.
The invention relates to a spatially-addressable set of single exon
nucleic acid probes for measuring gene expression in a sample derived
from human lung comprising single exon nucleic acid probes having one of
12614 nucleic acid sequences mentioned in the specification, or their
complements or the 12387 open reading frames derived from the 12614
probes. Also included are a microarray comprising the novel set of
probes; the novel set of probes which hybridize at high stringency to a
nucleic acid expressed in the human lung; measuring gene expression in a
sample derived from human lung, comprising (a) contacting the array with
a collection of detectably labeled nucleic acids derived from human lung
mRNA, and (b) measuring the label detectably bound to each probe of
the array; identifying exons in a eukaryotic genome, comprising
(a) algorithmically predicting at least one exon from genomic sequences
of the eukaryote; and (b) detecting specific hybridization of detectably
labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
having a fragment identical to the predicted exon, the probe is included
in the above mentioned microarray; assigning exons to a single gene,
comprising (a) identifying exons from genomic sequence by the method
above and (b) measuring the expression of each of the exons in several
tissues and/or cell types using hybridization to a single exon
microarrays having a probe with the exon, where a common pattern of
expression of the exons in the tissues and/or cell types indicates that
the exons should be assigned to a single gene; a peptide comprising one
of 12011 sequences, mentioned in the specification, or encoded by the
probes/open reading frames (ORF). The probes are used for gene
expression analysis, and for identifying exons in a gene, particularly
using human lung derived mRNA and for the study of lung diseases
such as asthma, lung cancer, chronic obstructive pulmonary disease
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
Niemann-Pick disease, Hernandez-Pudlak syndrome, sarcoidosis, pulmonary
haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
and hyaline membrane disease. The present sequence is a peptide/protein
encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp://wipo.int/pub/published_pct_sequences.
SQ Sequence 33 AA;
Alignment Scores:
Pred. No.: 3,67e-24 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservatave: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3,69% Indels: 0
DB: Gaps: 0
US-09-774-490-1 (1-2709) x ABG43013 (1-33)
QY 533 AAGAGATGCTAATTTCATCAGGTACTTAAGCATATATCAGACTCTTGACGCC 592
DB 1 LysglCysAlaAsnPhelIleValLeuYsaIaTyraSngGlnThrHisLeuTyAla 20
QY 593 TGTGGAACGGGGGCTTTTCATCCATTCGACCTACAT 631
DB 21 CysglYThrGlyAlaPheHisProIleCysThrTyIle 33
RESULT 37
AAV21292
ID AAV21292 standard; Protein; 31 AA.
XX
XX AAV21292;
AC
XX
XX 22-JUN-1999 (first entry)
DT
XX
XX Human semaphorin III mutant protein fragment 25.
DE
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
frameshift mutation; age-related disease; neurodegenerative disorder;
Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
Huntington's disease; multiple sclerosis; alcoholic liver disease;
diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
bc1-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
high mobility group protein-C; neuroendocrine specific protein A.
OS
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX MO9845322-A2.
XX
XX 15-OCT-1998.
PD
XX
XX 02-APR-1998; 98WO-1B00705.
PF
XX
XX 10-APR-1997; 97US-0043163.
PR
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI WPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type
 XX RNA
 XX
 PS Disclosure; Figure 16; 258bp; English.
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.
 CC
 SQ Sequence 31 AA;
 Alignment Scores:
 Pred. No.: 4,09e-22 Length: 31
 Score: 31.00 Matches: 31
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3,46% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21292 (1-31)
 QY 2190 CCTGGAGTCATTGACAGAGCATTTGGAGAACTCTTCAATAAGATGATGAG 2249
 DB 1 PGTTPlySErleuThrgInserleTTPlyAaenPheHelleYethetnetwctgu 20
 QY 2250 ATGGCTCTAAGACCAAGAAATGTCATAGCA 2282
 DB 21 MetAlaLeuArgProLySylScyProIleAla 31
 RESULT 38
 AAY21279 standard; Protein, 30 AA.
 ID AAY21279
 AC AAY21279;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 12.
 XX
 KW Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSF-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo.sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.

XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PI
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 PS Disclosure; Figure 16; 258bp; English.
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.
 CC
 SQ Sequence 30 AA;
 Alignment Scores:
 Pred. No.: 4,33e-21 Length: 30
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3,35% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21279 (1-30)
 QY 1143 ACTTTAAGATCCCTAAATCCAGTGTATGAGAGTGTATAGCACTCCAGTACATT 1202
 DB 1 ThrleuLysIleLeuLysIleGlnLeuYrmetGluCysLeuArgLeuProValThrPhe 20
 QY 1203 TCAAGGATCAGCCGTGTGTATGATAGCA 1232
 DB 21 SerArgGpsInProCysValCysIleAla 30
 RESULT 39
 AAY21269
 ID AAY21269 standard; Protein, 27 AA.
 AC AAY21269;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 2.
 XX
 KW Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;

KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SO Sequence 27 AA;
 Alignment Scores:
 Pred. No.: 5,11e-18 Length: 27
 Score: 27.00 Matches: 27
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.02% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21269 (1-27)
 QY 213 CTAGAGTTGTCTGTCTTTCTGGGAGATTAATTACAGCAAGCAAACTATCAGAATG 272
 Db 1 LeuGlyLeuSerValPheSerGlyCyluTyrrYrLeuGlnGlnGlnGlnThrIleArgMet 20
 QY 273 GGAAGAACAAATGTGCCAAGGC 293
 Db 21 GlyArgThrMetCysGlnGly 27

RESULT 40
 AAY21296
 ID AAY21296 standard; Protein; 26 AA.
 AC AAY21296;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 29.
 XX
 KW Human, beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SO Sequence 26 AA;
 Alignment Scores:
 Pred. No.: 5.41e-17 Length: 26

PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-Ap), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

CC Sequence 24 AA;

Alignment Scores:

Pred. No.:	6,05e-15	Length:	24
Score:	24.00	Matches:	24
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.68%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21285 (1-24)

QY 1509 ACCGATGAGTGCAGATGAGTATGATGTTATCGGAACAGATGTTGGA 1568

DB 1 ThrgutrpmeGlnLysmeCaspsermeCysLeuSerGlnGlnMetLeuGly 20

QY 1569 CCGTCTTAAG 1580

DB 21 Proheleuys 24

RESULT 43

ID AAY21337 standard; Protein; 23 AA.

XX AAY21337;

XX 22-JUL-1999 (first entry)

DT Human semaphorin III mutant protein fragment 70.

XX Human; beta-amyloid precursor protein; beta-Ap; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSF-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-1B00705.
 XX 10-APR-1997; 97US-0043163.

PA (UYUT-) RIJCKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-Ap), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

CC Sequence 23 AA;

Alignment Scores:

Pred. No.:	6.4e-14	Length:	23
Score:	23.00	Matches:	23
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.57%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21337 (1-23)

QY 2113 TCTACACAGAGATTCGCAATTACCTGCGATGCGTGAACATGCGTTGATCA 2172

DB 1 SerThrThrgLgUgLyPhaArgGlnLeuProCysGlyGlyThrtVAlHisThr 20

QY 2173 AACTTTCT 2181

DB 21 AsnSerSer 23

RESULT 44

ID AAY21310 standard; Protein; 23 AA.

XX AAY21310;

XX 22-JUL-1999 (first entry)

DT Human semaphorin III mutant protein fragment 43.

XX Human; beta-amyloid precursor protein; beta-Ap; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;

KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSGUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SQ Sequence 23 AA;
 Alignment Scores:
 Pred. No.: 6, 4e-14 Length: 23
 Score: 23.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2, 57% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY23110 (1-23)
 OY 778 TTTTATGGGGGAGAGCTTGTATCTTCGGAACCTGTTGGGACCAACCAATCAGAGAC 837
 DB 1 PheTYTGlyAlaArgLeucyStryIleuProAsnSerTrpAlaProProAsnGlnAsp 20
 OY 838 AGAGCAGCA 846

DB 21 ArgAlaAla 23
 RESULT 45
 AAY21290.
 ID AAY21290 standard; Protein; 23 AA.
 AC AAY21290;
 XX
 XX 22-JUL-1999 (first entry)
 DE Human semaphorin III mutant protein fragment 23.
 XX
 KW Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related diseases; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSGUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SQ Sequence 23 AA;

CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 67 AA;

XX Alignment Scores:

XX Pred. No.: 5.86e-13 Length: 67
XX Score: 22.00 Matches: 35
XX Percent Similarity: 97.22% Conservative: 0
XX Best Local Similarity: 97.22% Mismatches: 0
XX Query Match: 2.46% Indels: 1
XX DB: 21 Gaps: 0

US-09-774-490-1 (1-2709) x AAG02551 (1-67)

QY 176 AAGAGGACCTACGCGCTTCGACGATGGCTGTACTAGATTCGTCCTTTCTCG 235

DB 33 LysGlyThrIleValGlnArgLeuHisAlaHisGlnProPheGlnSerGlnHisasp 52

QY 236 GGAGTATTACTTACAGCAAGCAACTATCAGATGGAGACA 281

DB 53 -GluTyrTyrLeuGlnGlnGlnGlnThrIleArgMetGlyArgThr 67

RESULT 48

XX ID AAY21342 standard; Protein; 21 AA.

XX AAY21342;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 75.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX W09845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-1B00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.

XX (UTRO-) UNIV ROTTERDAM ERASMUS.

XX Burdach JPH, Grosveld FG, Van Leeuwen FW;

XX MPI; 1998-609901/51.

XX DR N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA

XX Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ Sequence 21 AA;

XX Alignment Scores:

XX Pred. No.: 7.17e-12 Length: 21
XX Score: 21.00 Matches: 21
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 2.35% Indels: 0
XX DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21342 (1-21)

QY 2293 CCAGAGGCTGTGTTACAGAGACTTCATGACCTCATCAACCCCAATCTCAACAGCAT 2352

DB 1 ProGluGlyLeuValGlnArgLeuHisAlaHisGlnProPheGlnSerGlnHisasp 20

QY 2353 GGA 2355

DB 21 Gly 21

RESULT 49

XX ID AAY21343 standard; Protein; 21 AA.

XX AAY21343;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 76.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX W09845322-A2.

XX 15-OCT-1998.

Db 21 Ser 21

RESULT 51

AAV21329

ID AAV21329 standard; Protein; 20 AA.

AC AAV21329;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 62.

Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; frameshift mutation; age-related disease; neurodegenerative disorder; Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; Huntington's disease; multiple sclerosis; alcoholic liver disease; diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; neurofilament-F; presentin I; presentin II; cellular tumour antigen; glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1; bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A; high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

XX Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UUKO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX MPI; 1998-609901/51.

XX N-PSDB; AAX75767.

Diagnosing disease by detecting frameshift mutations in RNA or corresponding protein mutations - used to diagnose cancer and neurological diseases, particularly Alzheimer's disease, and also for treatment and prevention with specific ribozymes or wild-type RNA

XX Disclosure; Figure 16; 258pp; English.

This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the use of neuronal system RNA molecules, specifically proteins including beta-amyloid precursor protein (beta-APP), the microtubule associated proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule associated protein 2 (MAP2), neurofilament-L, neurofilament-M, neurofilament-F, presentin I, presentin II, glial fibrillary acidic protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 20 AA;

Alignment Scores:

Pred. No.:	7,6e-11	Length:	20
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.23%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) * AAV21329 (1-20)

QY

1690 GGAGCAGACACTATATATGTTCAACGGCTGGGGTGGCCAGCTTCCTTACACCGGTG 1749

Db

1 AAlaAlaThrThrIleTyrIlePheAsnGlyTyrGlyCysProAlaPhePheThrProVal 20

RESULT 52

AAV21318

ID AAV21318 standard; Protein; 19 AA.

AC AAV21318;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 51.

Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; frameshift mutation; age-related disease; neurodegenerative disorder; Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; Huntington's disease; multiple sclerosis; alcoholic liver disease; diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; neurofilament-F; presentin I; presentin II; cellular tumour antigen; glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1; bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A; high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

XX Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UUKO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX MPI; 1998-609901/51.

XX N-PSDB; AAX75767.

Diagnosing disease by detecting frameshift mutations in RNA or corresponding protein mutations - used to diagnose cancer and neurological diseases, particularly Alzheimer's disease, and also for treatment and prevention with specific ribozymes or wild-type RNA

XX Disclosure; Figure 16; 258pp; English.

This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the use of neuronal system RNA molecules, specifically proteins including

CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ Sequence 19 AA;

Alignment Scores:

Pred. No.: 8.05e-10 Length: 19
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.12% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21304 (1-19)

QY 1045 ATGACATCTTCCTCAAGCTGCTGATTTGCTCAGNGCCAGGTCGAATGGCAT 1101

DB 1 MetaphenilleProGInSerSerSerAspLeuSerAlaArgSerLysTriPhis 19

RESULT 53

AAY21304 standard; Protein; 19 AA.

XX AAY21304;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 37.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UTRO-) UNIV ROTTERDAM ERASMUS.

XX Butbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI, 1998-609901/51.

XX N-P8DB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

XX corresponding protein mutations - used to diagnose cancer and

XX neurological diseases, particularly Alzheimer's disease, and also

XX for treatment and prevention with specific ribozymes or wild-type

XX RNA

XX Disclosure, Figure 16, 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX SQ Sequence 19 AA;

Alignment Scores:

Pred. No.: 8.05e-10 Length: 19
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.12% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21304 (1-19)

QY 574 TCAGACTCATTGACGCTGTGACACGGGGCTTTTCATTCACATTTGCACCTACAT 630

DB 1 SerAspSerLeuValArgLeuTriPAsnGlyGlyPheSerSerAsnLeuHis 19

RESULT 54

AAY21309 standard; Protein; 18 AA.

XX AAY21309;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 42.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UTRO-) UNIV ROTTERDAM ERASMUS.

XX Butbach JPH, Grosveld FG, Van Leeuwen FW;

XX

DR MPI; 1998-609901/51.
N-P8DB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
corresponding protein mutations - used to diagnose cancer and
neurological diseases, particularly Alzheimer's disease, and also
for treatment and prevention with specific ribozymes or wild-type
RNA

PS Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
caused by, or associated with, an RNA molecule that has a frameshift
mutation. The method is used to diagnose age-related diseases, especially
cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
and many others listed) or susceptibility to these disorders. The method
allows a definitive diagnosis of Alzheimer's disease in living patients,
at an early stage. It is based on the observation that disease may be
caused by mutations in RNA rather than DNA. The invention describes the
use of neuronal system RNA molecules, specifically proteins including
beta-amyloid precursor protein (beta-APP), the microtubule associated
proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
neurofilament-F, prenenilin I, prenenilin II, glial fibrillary acidic
protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
protein-C (HMGP-C) and neuroendocrine specific protein A.

SQ Sequence 18 AA;

Alignment Scores:
Pred. No.: 8.53e-09 Length: 18
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.01% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x AAV21309 (1-18)

Cy 721 GGTCGTGACAGCATCCCTTTAATGATGAGAATAATTACTCGGAACGGCACC 774
Db 1 AlaAlaapSerIleProPheasnArgTPARglieIleLeuuntpansCySer 18

RESULT 55
AAV21301
ID AAV21301 standard; Protein; 17 AA.

XX AAV21301;
XX
XX 22-JUL-1999 (first entry)
XX
XX Human semaphorin III mutant protein fragment 34.

Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
frameshift mutation; age-related disease; neurodegenerative disorder;
Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
Huntington's disease; multiple sclerosis; alcoholic liver disease;
diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
neurofilament-F; prenenilin I; prenenilin II; cellular tumour antigen;
glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.
OS Homo sapiens.
XX W09845322-A2.
XX
XX 15-OCT-1998.

PF	02-APR-1998;	98NW-IB00705.
XX		
PR	10-APR-1997;	97US-0043163.
XX		
PA	(UYUT-) RIJCSUNIV UTRECHT.	
XX	(ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.	
PA	(UVR0-) UNIV ROTTERDAM ERASMUS.	
XX		
P1	Burbach JPH, Grosveld FG, Van Leeuwen FW;	
XX		
DR	WPI, 1998-609901/51.	
XX	N-PSDB; AAX75767.	
PT	Diagnosing disease by detecting frameshift mutations in RNA or	
XX	corresponding protein mutations - used to diagnose cancer and	
PT	neurological diseases, particularly Alzheimer's disease, and also	
XX	for treatment and prevention with specific ribozymes or wild-type	
PT	RNA	
XX		
PS	Disclosure; Figure 16; 258pp; English.	
XX		
CC	This invention describes a novel method for the diagnosis of a disease	
XX	caused by, or associated with, an RNA molecule that has a frameshift	
CC	mutation. The method is used to diagnose age-related diseases, especially	
XX	cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's	
CC	disease, Down's syndrome, myotonic dystrophy, Huntington's disease, I	
XX	multiple sclerosis, alcoholic liver disease, diabetes mellitus type II	
CC	and many others listed) or susceptibility to these disorders. The method	
XX	allows a definitive diagnosis of Alzheimer's disease in living patients,	
CC	at an early stage. It is based on the observation that disease may be	
XX	caused by mutations in RNA rather than DNA. The invention describes the	
CC	use of neuronal system RNA molecules, specifically proteins including	
XX	beta-amyloid precursor protein (beta-ApP), the microtubule associated	
CC	proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule	
XX	associated protein 2 (Map2), neurofilament-L, neurofilament-M,	
CC	neurofilament-F, presentin I, presentin II, glial fibrillary acidic	
XX	protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma	
CC	2 (bc1-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group	
XX	protein-C (HMGP-C) and neuroendocrine specific protein A.	
SQ	Sequence	17 AA;
	Alignment Scores:	
	Pred. No.:	9.04e-08 Length: 17
	Score:	17.00 Matches: 17
	Percent Similarity:	100.00% Conservative: 0
	Best local Similarity:	100.00% Mismatches: 0
	Query Match:	1.90% Indels: 0
	DB:	19 Gaps: 0
US-09-774-490-1 (1-2709) x AAY21301 (1-17)		
Oy	451 TATCAAGATTTCCTTTCAAAGATTGTGCGCAGTAGTACTTACACGAGAAGA 501	
Dd		
	1 TygHnGlyPheserLysahpCyValAlaserlleuhHsgLnlySaNg 17	
RESULT 56		
ID	AAG62702 standard; peptide; 17 AA.	
XX	AAG62702	
XX	AAG62702;	
DT	17-SEP-2001 (first entry)	
XX		
DE	Hanatoxin-like sequence from type 3 semaphorin Sema3a.	
XX		
KM	Hanatoxin; tarantula; toxin; voltage-gated potassium channel;	
XX	voltage-gated calcium channel; hanatoxin-like sequence; HTUS;	
XX	semaphorin; dorsal root ganglion repulsion; growth cone collapse.	
OS	Mammalia.	
XX		

PN WO200138491-A2.
 XX 31-MAY-2001.
 XX 07-NOV-2000; 2000WO-US41943.
 XX 08-NOV-1999; 99US-0164056.
 XX (GENO) GEN HOSPITAL CORP.
 XX Behar O, Woolf CJ;
 XX WPI; 2001-451494/48.
 DR
 XX
 XX Polypeptide sequences that encompass the hanatoxin-like sequences of
 PT semaphorin, useful as a drugs to treat any condition or disease that
 PT is characterized by abnormal calcium channel function -
 XX
 XX Claim 3; Page 10; 29pp; English.
 XX
 XX AAG62702-22 represent hanatoxin-like sequences (HTLS) found in
 CC the semaphorin domain of mammalian secreted semaphorins. Hanatoxin
 CC is a tarantula toxin that selectively blocks some voltage-gated
 CC potassium and calcium channels. The HTLS is responsible for the dorsal
 CC root ganglion repulsion and growth cone collapse activities associated
 CC with semaphorins. Polypeptides containing HTLS can be used to modulate
 CC the activity of calcium channels. The peptides can also be used as an
 CC antigen to generate antibodies that can then be used to modulate the
 CC activity of calcium channels by inactivating naturally occurring channel
 CC ligands. The peptides or antibodies can be used as drugs to treat any
 CC condition or disease that is characterized by abnormal calcium channel
 CC function.
 CC
 SO Sequence 17 AA:
 Alignment Scores:
 Pred. No.: 9,04e-08 Length: 17
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x AAG62702 (1-17)
 QY 1763 AAGGTCGTGAGTGGCTGCGCCGAGACCTTACTGCTTGGAT 1813
 Db 1 LysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIrrPasp 17
 RESULT 57
 AAB84219
 ID AAB84219 standard; Protein; 779 AA.
 AC AAB84219;
 XX
 XX 06-AUG-2001 (first entry)
 DT
 XX
 XX Amino acid sequence of a semaphorin polypeptide designated ZSMF-16.
 DE
 XX
 XX Human; semaphorin; ZSMF-16; neurite growth; neurite outgrowth;
 KW T lymphocyte suppressor; cancer; stroke; brain damage; paralysis;
 KW spinal injury; neurodegenerative disease; amyotrophic lateral sclerosis;
 KW Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW peripheral neuropathy; demyelinating disease; multiple sclerosis;
 KW immunosuppression; autoimmune disease; insulin dependent diabetes;
 KW rheumatoid arthritis.
 KW
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..22
 FT /note= "signal peptide"
 FT Protein 23..779

PT /note= "mature protein"
 PT Modified-site 62
 FT /note= "N-glycosylation site"
 FT Domain 76..500
 FT /note= "semaphorin"
 FT Modified-site 124
 FT /note= "N-glycosylation site"
 FT Domain 593..654
 FT /note= "Ig-like domain"
 FT Modified-site 594
 FT /note= "N-glycosylation site"
 PN WO200140278-A2.
 XX
 XX 07-JUN-2001.
 PD
 XX
 XX 06-DEC-2000; 2000WO-US31116.
 PF
 XX
 XX 06-DEC-1999; 99US-0455560.
 PR
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX
 XX Holloway JL, Foley KP;
 PI
 XX
 XX WPI; 2001-374784/39.
 DR N-PSDB; AAF90250.
 XX
 XX Novel human semaphorin polypeptide, ZSMF-16, useful for treating
 PT peripheral neuropathies Alzheimer's and Huntington's disease and
 PT polynucleotide encoding ZSMF-16 useful for detecting genetic
 PT abnormality and cancer -
 XX
 XX
 PS Claim 13; Page 121-123; 124pp; English.
 XX
 XX The present sequence represents a semaphorin polypeptide, designated
 CC ZSMF-16. ZSMF-16 is a neurite growth and development modulator. It also
 CC enhances spinal cord and sensory neurite outgrowth and patterning, and
 CC is involved in the activation and regulation of T lymphocytes suppressor.
 CC ZSMF-16 is useful for detecting a genetic abnormality or cancer. ZSMF-16
 CC polynucleotide probes can be used to detect 3p21 loss, trisomy,
 CC duplication or translocation associated with mammary tumor tissue,
 CC breast tumour, liver, small intestine, bone cancers, etc.. ZSMF-16 can
 CC be used to modulate neurite growth and development and demarcate nervous
 CC system structures. ZSMF-16 are also useful for regenerating and directing
 CC neurite outgrowths following strokes, brain damage caused by head
 CC injuries, paralysis caused by spinal injuries, and for treating
 CC neurodegenerative diseases such as amyotrophic lateral sclerosis,
 CC Alzheimer's disease, Huntington's disease, Parkinson's disease and
 CC peripheral neuropathies, or demyelinating diseases e.g., multiple
 CC sclerosis. ZSMF-16 also acts as a mediator of immunosuppression,
 CC and thus useful for diagnosing and treating autoimmune diseases such as
 CC insulin dependent diabetes, rheumatoid arthritis, and multiple
 CC sclerosis. It can also be used as an anti-inflammatory for inhibition
 CC of antigen in humoral and cellular immunity and for immunosuppression
 CC in graft and organ transplants.
 CC
 SO Sequence 779 AA:
 Alignment Scores:
 Pred. No.: 5.51e-08 Length: 779
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x AAB84219 (1-779)
 QY 1766 GCGTGTGCTGAGTGGCTGCGCCGAGACCTTACTGCTTGGAT 1816
 Db 525 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIrrPaspGly 541
 RESULT 58

AA678481
 ID AAG78481 standard; Protein; 779 AA.
 XX
 AC AAG78481;
 XX
 DT 12-APR-2002 (first entry)
 XX
 DE Human ZSMF-16.
 XX
 KW Semaphorin; cancer; immune response; ZSMF-16; human;
 KW neuroprotective; nootropic; multiple sclerosis;
 KW Alzheimer's disease; Parkinson's disease; stroke.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22 /label= signal_peptide
 FT Protein 23..779 /label= mature_protein
 FT Region 23..75 /note= "N-terminal region"
 FT Modified-site 62 /note= "N- glycosylated"
 FT Domain 76..500 /label= semaphorin_domain
 FT Modified-site 124 /note= "N- glycosylated"
 FT Domain 501..592 /label= Middle_domain
 FT Domain 593..654 /label= Ig-like_domain
 FT Modified-site 594 /note= "N- glycosylated"
 FT Domain 655..738 /label= C-terminal_domain
 FT Region 744..751 /note= "ATP/GTP binding site motif"
 XX
 XX
 PN US2001049432-A1.
 XX
 PD 06-DEC-2001.
 XX
 PF 06-DEC-2000; 2000US-0731179.
 XX
 PR 06-DEC-1999; 99US-169238P.
 XX
 PA (HOLL/) HOLLOWAY J L.
 PA (FOLE/) FOLEY K P.
 XX
 PI Holloway JL, Foley KP;
 XX
 DR WPI; 2002-121401/16.
 DR N-PSDB; ABA97362.
 XX
 XX
 PT Polynucleotides encoding a semaphorin polypeptide expressed in breast
 PT cancer and neuronal tissue, useful for treating neurological damage and
 PT disease, and to detect cancer -
 XX
 PS Claim 1; Page 33-35; 59pp; English.
 XX
 CC This invention relates to an isolated polynucleotide encoding a
 CC semaphorin polypeptide expressed in breast cancer and neuronal tissue.
 CC The polypeptide is neuroprotective and nootropic in its action.
 CC The semaphorin polypeptide can be used to modulate neurite growth
 CC and development, for example to treat peripheral neuropathies, stroke,
 CC brain damage from brain and spine injuries, and neurodegenerative
 CC diseases such as multiple sclerosis, Alzheimer's and Parkinson's
 CC disease. It can also be used to modulate cellular activation,
 CC targeting, adhesion, proliferation and differentiation and
 CC immunological responses. Molecules of the invention can also be used to
 CC detect cancer. This sequence represents the ZSMF-16 which is a member of
 CC the semaphorin family.

XX
 SQ Sequence 779 AA;
 Alignment Scores: 5.51e-08 Length: 779
 Pred. No.: 17.00 Matches: 17
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.90% Gaps: 0
 DB: 23
 US-09-774-490-1 (1-2709) x AAG78481 (1-779)
 QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGAT 1816
 |||||
 Db 525 AAcYsAlaGluCysCysLeuAlaArgAspProTyrCysAlaIatPAspGly 541
 RESULT 59
 AAB23609
 ID AAB23609 standard; Protein; 782 AA.
 XX
 AC AAB23609;
 XX
 DT 12-JAN-2001 (first entry)
 XX
 DE Human secreted protein SEQ ID NO: 18.
 XX
 KW Human; secreted protein; cytokine; cell proliferation;
 KW nutritional supplement; immune modulation; autoimmune disorder;
 KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 10..22 /label= signal_peptide
 FT Protein 23..782 /label= mature_protein
 XX
 PN WO200049134-A1.
 XX
 PD 24-AUG-2000.
 XX
 PF 18-FEB-2000; 2000WO-US04340.
 XX
 PR 19-FEB-1999; 99US-0120680.
 PR 23-APR-1999; 99US-0296733.
 PR 17-AUG-1999; 99US-0149639.
 PR 23-SEP-1999; 99US-0155686.
 PR 01-OCT-1999; 99US-0157247.
 PR 29-NOV-1999; 99US-0167823.
 PR 29-NOV-1999; 99US-0167823.
 PR 15-FEB-2000; 2000US-0296733.
 XX
 PA (ALPH-) ALPHAGEN INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 XX
 DR WPI; 2000-549267/50.
 DR N-PSDB; AAA93109.
 XX
 PT New secreted proteins and polynucleotides encoding them, which are
 PT derived from Homo sapiens, useful for therapy, diagnosis, and research,
 PT as well as nutritional sources or supplements -
 XX
 PS Claim 27; Page 250-253; 309pp; English.
 XX
 CC The present sequence is the sequence of a human secreted protein. Its
 CC cDNA was isolated from an adult brain cDNA library. The proteins
 CC and coding sequences of the invention can be used in the isolation of
 CC similar genes and proteins, in the elucidation of their function in vivo,
 CC and to treat a number of conditions. It is possible that they may have
 CC uses as nutritional supplements, as cytokine or cell proliferation

CC factors, in immune modulation, where they may be used to treat immune and
CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
CC have chemokine or chemotactic activity, haemostatic or thrombolytic
CC activity, or anti-inflammatory activity.

XX Sequence 782 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	782
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB23609 (1-782)

OY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGAGTGT 1816

DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIleTpaSpGly 544

RESULT 60

ID AAG65620 standard; Protein; 782 AA.

XX AAG65620;

DT 07-JAN-2002 (first entry)

DE Novel human protein (NHP) sequence.

KW NHP; novel human protein; secreted protein; semaphorin; oxytocin;

KW neurohypophyseal; neotropic; gene therapy; drug screening.

XX Homo sapiens.

XX WO200170806-A2.

XX PD 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US088834.

XX 20-MAR-2000; 2000US-190638P.

XX 22-MAR-2000; 2000US-191188P.

XX 31-MAR-2000; 2000US-193639P.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G;

XX Scoville J;

XX WPI, 2001-611483/70.

XX N-PSDB; AA47791.

XX New polynucleotides encoding human proteins that share structural

XX similarity with semaphorin proteins, protein hormones of

XX neurohypophyseal family for drug screening, diagnosis and therapy of

XX biological disorders

XX Claim 4; Page 38-40; 43pp; English.

XX The invention relates to novel human secreted proteins (NHP) that share
XX structural similarity with semaphorin proteins, protein/peptide hormones
XX of the neurohypophyseal family and oxytocin (neurohypophysis I precursor)
XX family. The NHP nucleotide sequences are useful in drug screening of
XX techniques for treating symptomatic or phenotypic manifestations of
XX perturbing the normal function of NHP in the body. Nucleotide constructs
XX encoding NHP products are useful in gene therapy for modulating NHP
XX expression. The constructs can be used to genetically engineer host cells
XX to express NHP products in vivo, these genetically engineered cells
XX function as bioreactors in the body delivering a continuous supply of a
XX NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences

CC also find use in molecular mutagenesis/evolution of proteins that are
CC partially encoded by the NHP sequences. The encoded NHP polypeptides are
CC useful for generating antibodies, as reagents in diagnostic assays, for
CC identifying other cellular gene products related to NHP and as reagents
CC in assays for screening for compounds that are useful in the treatment of
CC mental, biological or medical disorders and diseases. The present
CC sequence represents the amino acid sequence of a NHP.

XX Sequence 782 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	782
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG65620 (1-782)

OY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGAGTGT 1816

DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIleTpaSpGly 544

RESULT 61

ID AAB23636 standard; Protein; 785 AA.

XX AAB23636;

DT 12-JAN-2001 (first entry)

DE Human secreted protein SEQ ID NO: 92.

KW Human secreted protein; cytokine; cell proliferation;

KW nutritional supplement; immune modulation; autoimmune disorder;

KW haematopoiesis regulation; tissue growth; haemostasis; inflammation.

XX Homo sapiens.

XX WO200049134-A1.

XX PD 24-AUG-2000.

XX 18-FEB-2000; 2000WO-US04340.

XX 19-FEB-1999; 99US-0120680.

XX 23-APR-1999; 99US-0298733.

XX 17-AUG-1999; 99US-0149639.

XX 23-SEP-1999; 99US-0155686.

XX 01-OCT-1999; 99US-0157247.

XX 29-NOV-1999; 99US-0167822.

XX 29-NOV-1999; 99US-0167823.

XX 15-FEB-2000; 2000US-0298733.

XX (ALPH-) ALPHAGEN INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX WPI, 2000-549267/50.

XX New secreted proteins and polynucleotides encoding them, which are
XX derived from Homo sapiens, useful for therapy, diagnosis, and research,
XX as well as nutritional sources or supplements -
XX Disclosure; Page 298-300; 309pp; English.
XX The present invention is concerned with a number of secreted proteins
XX and their coding sequences isolated from various human cDNA libraries.
XX The proteins and coding sequences can be used in the isolation of
XX similar genes and proteins, in the elucidation of their function in vivo,
XX and to treat a number of conditions. It is possible that they may have
XX uses as nutritional supplements, as cytokine or cell proliferation

CC factors, in immune modulation, where they may be used to treat immune and
CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
CC have chemokine or chemotactic activity, haemostatic or thrombolytic
CC activity, or anti-inflammatory activity. No information about sequences
CC AAB2632-B2645 is given in the specification.

XX SQ Sequence 785 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	785
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB2636 (1-785)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816

Db 531 AAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 547

RESULT 62

AAG65619

ID AAG65619 standard; Protein; 875 AA.

AC AAG65619;

DT 07-JAN-2002 (first entry)

DE Novel human protein (NHP) sequence.

KW NHP; novel human protein; secreted protein; semaphorin; oxytocin;

KM neurohypophyseal; neurotropic; gene therapy; drug screening.

OS Homo sapiens.

PN WO200170806-A2.

PD 27-SEP-2001.

PF 20-MAR-2001; 2001WO-US08834.

PR 20-MAR-2000; 2000US-190638P.

PR 22-MAR-2000; 2000US-191188P.

PR 31-MAR-2000; 2000US-193639P.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G;

PI Scoville J;

DR WPI; 2001-611483/70.

DR N-PSDB; AA47790.

PT New polynucleotides encoding human proteins that share structural

PT similarity with semaphorin proteins, protein hormones of

PT neurohypophyseal family for drug screening, diagnosis and therapy of

PT biological disorders -

PS Claim 2; Page 35-37; 43pp; English.

XX The invention relates to novel human secreted proteins (NHP) that share

XX structural similarity with semaphorin proteins, protein/peptide hormones

XX of the neurohypophyseal family and oxytocin (neurophysin 1 precursor)

XX family. The NHP nucleotide sequences are useful in drug screening

XX techniques for treating symptomatic or phenotypic manifestations of

XX perturbing the normal function of NHP in the body. Nucleotide constructs

XX encoding NHP products are useful in gene therapy for modulating NHP

XX expression. The constructs can be used to genetically engineer host cells

XX to express NHP products in vivo, these genetically engineered cells

XX function as bioreactors in the body delivering a continuous supply of a

CC NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences
CC also find use in molecular mutagenesis/evolution of proteins that are
CC partially encoded by the NHP sequences. The encoded NHP polypeptides are
CC useful for generating antibodies, as reagents in diagnostic assays, for
CC identifying other cellular gene products related to NHP and as reagents
CC in assays for screening for compounds that are useful in the treatment of
CC mental, biological or medical disorders and diseases. The present
CC sequence represents the amino acid sequence of a NHP.

XX SQ Sequence 875 AA;

Alignment Scores:

Pred. No.:	5.42e-08	Length:	875
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG65619 (1-875)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816

Db 621 AAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 637

RESULT 63

AAG62704

ID AAG62704 standard; peptide; 17 AA.

AC AAG62704;

DT 17-SEP-2001 (first entry)

DE Hanatoxin-like sequence from type 3 semaphorin msemab3B.

KW Hanatoxin; tarantula; toxin; voltage-gated potassium channel;

KM voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.

OS Mus sp.

PN WO200138491-A2.

PD 31-MAY-2001.

PR 07-NOV-2000; 2000WO-US41943.

PR 08-NOV-1999; 99US-0164056.

PA (GENO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

PI WPI; 2001-451494/48.

PT Polypeptide sequences that encompass the hanatoxin-like sequences of

PT semaphorins, useful as a drug to treat any condition or disease that

PT is characterized by abnormal calcium channel function -

PS Claim 3; Page 10; 29pp; English.

XX AAG62702-22 represent hanatoxin-like sequences (HTLS) found in

XX the semaphorin domain of mammalian secreted semaphorins. Hanatoxin

XX is a tarantula toxin that selectively blocks some voltage-gated

XX potassium and calcium channels. The HTLS is responsible for the dorsal

XX root ganglion repulsion and growth cone collapse activities associated

XX with semaphorins. Polypeptides containing HTLS can be used to modulate

XX the activity of calcium channels. The peptides can also be used as an

XX antigen to generate antibodies that can then be used to modulate the

XX activity of calcium channels by inactivating naturally occurring channel

XX ligands. The peptides or antibodies can be used as drugs to treat any

XX condition or disease that is characterized by abnormal calcium channel

XX function.

XX SQ Sequence 17 AA;
 Alignment Scores:
 Pred. No.: 9.52e-07 Length: 17
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.79% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x AAG62704 (1-17)
 OY 1766 GCGTGTGCTGAGTGTGCTCGCCGACACCTTACTGCTGGAT 1813
 DB 2 AAcGSAAGlUCGcGcSleuAlaArgaProTyrGSAItTPhsp 17
 RESULT 64
 ID AAY21300 standard; Protein; 15 AA.
 AC AAY21300;
 XX 22-JUL-1999 (first entry)
 DT Human semaphorin III mutant protein fragment 33.
 DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 XX frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACADEMIES & SCI.
 PA (UTRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Butbach JPH, Grosvelde FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure: Figure 16, 258bp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method

CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 15 AA;
 Alignment Scores:
 Pred. No.: 1.02e-05 Length: 15
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21300 (1-15)
 OY 403 GCTGTATGTTGACGAAAGATCACATTTTCATTCGACCTGGT 447
 DB 1 AAlaValCysTrpSerIleGclSerHisIleHehIleArgProGly 15
 RESULT 65
 ID AAY21281 standard; Protein; 15 AA.
 AC AAY21281;
 XX 22-JUL-1999 (first entry)
 DT Human semaphorin III mutant protein fragment 14.
 DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 XX frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACADEMIES & SCI.
 PA (UTRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Butbach JPH, Grosvelde FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type

PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HMPF-1, high mobility group
 CC protein-C (HMG-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 15 AA;

Alignment Scores:
 Pred. No.: 1,02e-05 Length: 15
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21281 (1-15)

QY 1401 CCTTGCAAGAGTCATCCAGCCATGACATCCAGTGTTCCTA 1445

DB 1 ProluGIngluValIleGInProCysThrIleGInCysPheIeu 15

RESULT 66

AA662727 standard; peptide; 749 AA.

AC AA662727;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of human semaphorin Sema3B.

KW Hanatoxin; tarentula; toxin; voltage-gated potassium channel;

KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.

OS Homo sapiens.

PN WO200138491-A2.

PD 31-MAY-2001.

PF 07-NOV-2000; 2000WO-US41943.

PR 08-NOV-1999; 99US-0164056.

PA (GENO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

DR WPI; 2001-451494/48.

PT Polypeptide sequences that encompass the hanatoxin-like sequences of

PT semaphorin, useful as a drugs to treat any condition or disease that

PT is characterized by abnormal calcium channel function -

PS Claim 6; Page 12; 29pp; English.

XX
 CC The present sequence represents a semaphorin. The specification
 CC describes hanatoxin-like sequences (HTLS) found in the semaphorin
 CC domain of mammalian secreted semaphorins. Hanatoxin is a tarentula
 CC toxin that selectively blocks some voltage-gated potassium and calcium
 CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
 CC and growth cone collapse activities associated with semaphorins.
 CC Polypeptides containing HTLS can be used to modulate the activity of
 CC calcium channels. The peptides can also be used as an antigen to
 CC generate antibodies that can then be used to modulate the activity of
 CC calcium channels by inactivating naturally occurring channel ligands.
 CC The peptides or antibodies can be used as drugs to treat any condition
 CC or disease that is characterized by abnormal calcium channel function.

SQ Sequence 749 AA;

Alignment Scores:
 Pred. No.: 6.13e-06 Length: 749
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AA662727 (1-749)

QY 1028 AGAAGTCGCGAATTAATGACACATTCCTCAAGCTGCTG 1072

DB 277 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 291

RESULT 67

AAW30617 standard; Protein; 751 AA.

AC AAW30617;

DT 31-MAR-1999 (first entry)

DE Human semaphorin E protein from clone BR5334.

KW Human; semaphorin E; clone BR5334; nutritional; immune stimulating;

KW vaccine; haematopoiesis regulating activity; tissue growth;

KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;

KW gene therapy.

OS Homo sapiens.

PN WO9853065-A1.

PD 26-NOV-1998.

PF 19-MAY-1998; 98WO-US10188.

PR 18-MAY-1998; 98US-0080695.

PR 19-MAY-1997; 97US-0858834.

PA (GENO) GENETICS INST INC.

PI Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavalie ER;

PI McCoy JM, Werberg D, Racie LA, Spaulding V, Treacy M;

DR WPI; 1999-059742/05.

DR N-PDB; AAX03792.

PT New polypeptides encoding secreted human proteins - derived from a

PT human foetal kidney cDNA library

PS Claim 7; Page 44-47; 58pp; English.

CC The present sequence is human semaphorin E from cDNA clone BR5334.

CC Human semaphorin E polynucleotide sequences and protein sequences from

CC the present invention, are predicted to have biological activities which

CC would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data is
CC given. Suggested activities include nutritional activity, immune
CC stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis
CC regulating activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, and tumour inhibition
CC activity. The polynucleotide sequences are also stated to be useful for
CC gene therapy. A host cell transfected with the polynucleotide sequence
CC for recombinant production of the clone BR5334 related protein.

SQ Sequence 751 AA;

Alignment Scores:

Pred. No.:	6.13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	20	Gaps:	0

US-09-774-490-1 (1-2709) x AAM30617 (1-751)

QY 1028 AGAAGCTGTGTAATTAATGACAACTTCCTCAAGCTCGTCTG 1072

DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 68

ID AAB28379 standard; Protein; 751 AA.

XX AAB28379;

XX 19-FEB-2001 (first entry)

XX Clone BR533_4.

XX Human; clone BR533_4; secreted protein; immune deficiency;
XX microbial infection; autoimmune disorder; allergy; asthma; inflammation;
XX myeloid deficiency; lymphoid cell deficiency; anaemia; burn;
XX wound healing; ulcer; periodontal disease; nervous system disease;
XX neuropathy; lung fibrosis; liver fibrosis; cancer.

XX Homo sapiens.

XX WO200063692-A1.

XX 26-OCT-2000.

XX 14-APR-2000; 2000MO-US10048.

XX 15-APR-1999; 99US-0292550.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy W, Spaulding V;

XX WPI; 2000-679620/66.

XX N-PSDB; AAC6800.

XX New monoclonal antibodies, useful for treating cancer and
XX immunodetection of secreted proteins which are in turn useful for
XX treating neurological, inflammatory, immune diseases and microbial
XX infections -
XX Claim 8; Pages 67-69; 75pp; English.

XX The present sequence is human clone BR533_4. The coding sequence for the
XX present sequence was isolated from a human foetal kidney cDNA library.
XX The BR533_4 protein is a secreted protein and can be used to treat a

CC number of conditions including various immune deficiencies and disorders,
CC microbial infections, autoimmune disorders, allergic reactions such as
CC asthma, respiratory problems, inflammation, myeloid or lymphoid cell
CC deficiencies, anemias, burns, wound healing, ulcers, periodontal
CC disease, central and peripheral nervous system diseases and neuropathies,
CC lung or liver fibrosis and cancer.

SQ Sequence 751 AA;

Alignment Scores:

Pred. No.:	6.13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB28379 (1-751)

QY 1028 AGAAGCTGTGTAATTAATGACAACTTCCTCAAGCTCGTCTG 1072

DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 69

ID AAG62728 standard; peptide; 751 AA.

XX AAG62728;

XX 17-SEP-2001 (first entry)

XX Amino acid sequence of human semaphorin Sema3C.

XX Hanatoxin; taramula; toxin; voltage-gated potassium channel;
XX voltage-gated calcium channel; hanatoxin-like sequence; HTLS;
XX semaphorin; dorsal root ganglion repulsion; growth cone collapse.

XX Homo sapiens.

XX WO200138491-A2.

XX 31-MAY-2001.

XX 07-NOV-2000; 2000MO-US41943.

XX 08-NOV-1999; 99US-0164056.

XX (GENO) GEN HOSPITAL CORP.

XX Behar O, Woolf CJ;

XX WPI; 2001-451494/48.

XX Polypeptide sequences that encompass the hanatoxin-like sequences of
XX semaphorins, useful as a drugs to treat any condition or disease that
XX is characterized by abnormal calcium channel function -
XX Claim 6; Page 12; 29pp; English.

XX The present sequence represents a semaphorin. The specification
XX describes hanatoxin-like sequences (HTLS) found in the semaphorin
XX domain of mammalian secreted semaphorins. Hanatoxin is a taramula
XX toxin that selectively blocks some voltage-gated potassium and calcium
XX channels. The HTLS is responsible for the dorsal root ganglion repulsion
XX and growth cone collapse activities associated with semaphorins.
XX Polypeptides containing HTLS can be used to modulate the activity of
XX calcium channels. The peptides can also be used as an antigen to
XX generate antibodies that can then be used to modulate the activity of
XX calcium channels by inactivating naturally occurring channel ligands.
XX The peptides or antibodies can be used as drugs to treat any condition
XX or disease that is characterized by abnormal calcium channel function.
XX Sequence 751 AA;

Alignment Scores:

Pred. No.: 6,13e-06 Length: 751
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAG62728 (1-751)

QY 1028 AGAAGTCTGTAATGAACAACATTCTCAAGCTGCTG 1072
 |||||
 DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 70

ABP68623 standard; Protein: 751 AA.

AC ABP68623;

DT 14-JAN-2003 (first entry)

DE Human pancreatic cancer expressed protein SEQ ID NO 172.

KM Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;

KW cytostatic; tumour.

OS Homo sapiens.

PN WO200260317-A2.

PD 08-AUG-2002.

PF 30-JAN-2002; 2002MO-US02781.

PR 30-JAN-2001; 2001US-265305P.

PR 31-JAN-2001; 2001US-265682P.

PR 09-FEB-2001; 2001US-267568P.

PR 21-MAR-2001; 2001US-278651P.

PR 28-APR-2001; 2001US-287112P.

PR 16-MAY-2001; 2001US-291631P.

PR 12-JUL-2001; 2001US-305484P.

PR 20-AUG-2001; 2001US-313999P.

PR 27-NOV-2001; 2001US-333626P.

XX (CORI-) CORIXA CORP.

PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;

DR WPI; 2002-627435/67.

DR N-PSDB; ABV94769.

XX New isolated polynucleotide and pancreatic tumor polypeptides, useful

PT for diagnosing, preventing and/or treating cancer, particularly

PT pancreatic cancer

XX Claim 2; SEQ ID NO 172; 300pp + Sequence Listing; English.

PS The invention relates to an isolated polynucleotide (I) comprising: (a)

XX any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);

CC (b) complements of (a); (c) sequences consisting of at least 20

CC contiguous residues of (a); (d) sequences that hybridize to (a), under

CC moderately stringent conditions; (e) sequences having at least 75% or 90%

CC identity to (a); or (f) degenerate variants of (a). Polypeptides

CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to

CC detect cancer in a patient and compositions comprising polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations and

CC antigen presenting cells expressing the polypeptide are useful in

CC treating pancreatic cancer and stimulating an immune response. The

CC polynucleotides can be used as probes or primers for nucleic acid

CC hybridisation, in the design and preparation of ribozyme molecules for

CC inhibiting expression of the tumour polypeptides and proteins in the

CC tumour cells, in vaccines and for gene therapy.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcc_sequences.

SO Sequence 751 AA;

Alignment Scores:

Pred. No.: 6,13e-06 Length: 751
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x ABP68623 (1-751)

QY 1028 AGAAGTCTGTAATGAACAACATTCTCAAGCTGCTG 1072
 |||||
 DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 71

AAV21315 standard; Protein: 14 AA.

AC AAV21315;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 48.

KM Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;

KW frameshift mutation; age-related disease; neurodegenerative disorder;

KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

KM Huntington's disease; multiple sclerosis; alcoholic liver disease;

KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;

KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;

KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;

KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;

KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSF-A;

KM high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

OS Homo sapiens.

PN WO9845322-A2.

PD 15-OCT-1998.

PD 02-APR-1998; 98WO-1B00705.

PF 10-APR-1997; 97US-0043163.

PR (UYUT-) RIJKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Butbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

DR N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type

PT RNA

PS Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease

CC caused by, or associated with, an RNA molecule that has a frameshift

CC mutation. The method is used to diagnose age-related diseases, especially

CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II CC and many others listed) or susceptibility to these disorders. The method CC allows a definitive diagnosis of Alzheimer's disease in living patients, CC at an early stage. It is based on the observation that disease may be CC caused by mutations in RNA rather than DNA. The invention describes the CC use of neuronal system RNA molecules, specifically proteins including CC beta-amyloid precursor protein (beta-APP), the microtubule associated CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M, CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group CC protein-C (HMGP-C) and neuroendocrine specific protein A.

CC
XX
SQ Sequence 14 AA;

Alignment Scores:

Pred. No.:	0.000108	Length:	14
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21276 (1-14)

QY 946 AATGCAATGATGAGAACACTCTGGAAAGCTACTACGC 987

Db 1 LysCysAsnArgTrpArgThrLeuTrpLysSerTySerArg 14

RESULT 72

AAY21276

ID AAY21276 standard; Protein; 14 AA.

AC AAY21276;

XX 22-JUL-1999 (first entry)

DT Human semaphorin III mutant protein fragment 9.

DE

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; KW frameshift mutation; age-related disease; neurodegenerative disorder; KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; KW Huntington's disease; multiple sclerosis; alcoholic liver disease; KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; KW neurofilament-F; presentin I; presentin II; cellular tumour antigen; KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1; KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A; KW high mobility group protein-C; neuroendocrine specific protein A.

XX
OS Synthetic.

OS Homo sapiens.

XX MO9845322-A2.

XX 15-OCT-1998.

PD 02-APR-1998; 98MO-IB00705.

XX 10-APR-1997; 97US-0043163.

PR (UYUT-) RIJKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UNRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

PI MPI: 1998-609901/51.

DR N-ESDB; AAX75767.

XX

PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA

XX
PS Disclosure; Figure 16; 258pp; English.

CC
XX
CC Caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.

CC
XX
SQ Sequence 14 AA;

Alignment Scores:

Pred. No.:	0.000108	Length:	14
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21276 (1-14)

QY 996 GTCGATATGCAAGATGACTTTGGAGGGCAGAGAGCTCG 1037

Db 1 ValArgTyrAlaArgMetThrLeuGlnGlyThrGluValTrp 14

RESULT 73

AAY27127

ID AAY27127 standard; Protein; 777 AA.

AC AAY27127;

XX 14-SEP-1999 (first entry)

DT Human brain tissue-derived polypeptide (clone OM007).

DE

XX Human tissue; human; bone marrow; umbilical cord venous endothelial cell; KW recombinant; diagnosis; treatment.

XX
OS Homo sapiens.

XX MO9933873-A1.

XX 08-JUL-1999.

PD 25-DEC-1998; 98MO-JP05952.

XX 26-DEC-1997; 97JP-0358811.

PR (ONOV) ONO PHARM CO LTD.

XX

XX PI Fukushima D, Shibayama S, Tada H;
XX
DR WPI: 1999-419088/35.
DR N-PSDB; AAX89112, AAX89113.
XX
PT New adult human brain tissue-produced polypeptides useful for
PT diagnosis and treatment
XX
PS Claim 1; Page 36-39; 86pp; Japanese.
XX
CC The invention provides polypeptides (AAV27127-Y27133) produced by human
CC adult brain tissue, human bone marrow or a human umbilical cord venous
CC endothelial cell. Host cells transformed with vectors comprising the
CC nucleic acids encoding the polypeptides are used for the recombinant
CC expression of the polypeptides. The polypeptides can be used in
CC diagnosis, treatment and basic studies, with wide applications in
CC treatment depending on the activity to be aimed at. Sequences
CC AAX89112-135 represent nucleic acids encoding the polypeptides.
XX
SQ Sequence 777 AA;

Alignment Scores:
Pred. No.: 6,42e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x AAY27127 (1-777)

QY 1040 AATAATGGACACATTCTCAAGCTCGTCGATTGCTCA 1081
Db |||||||||||||||||||||||||||||||||||||||
298 AenlystlrprhtrhrheueySalAAgveuileCysSer 311

RESULT 74
ID AAY99427
AAV99427 standard; Protein; 777 AA.
XX
AC AAY99427;
XX
DT 08-AUG-2000 (first entry)
XX
DE Human PRO191 (UNQ760) amino acid sequence SEQ ID NO:310.
XX
KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
XX
OS Homo sapiens.
XX
PN WO200012708-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WC-US20111.
XX
PR 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.

PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.

PT to screen for modulators of the compounds -
 XX
 PS
 XX Claim 11, Fig 348; 774pp; English.

CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.
 XX
 SQ Sequence 777 AA:

Alignment Scores:
 Pred. No.: 6.42e-05 Length: 777
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.56% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAU29197 (1-777)

QY 1040 AATAATGGACACATTCCCAAGCTCGTGTGATTTGCTCA 1081
 ||||||||||||||||||||||||||||||||||||||||
 Db 298 AsnLysIrpIrnTrpNeuLeuLysAlaArgLeuIleCysSer 311

Search completed: August 3, 2003, 10:49:42
 Job time : 116.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: August 3, 2003, 10:47:09 ; Search time 57 Seconds
(without alignments)
11288.429 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aatcttattatcagatg.....agcctttttcctaacc 2709

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 451899 segs, 118759770 residues

Word size: 1

Total number of hits satisfying chosen parameters: 847952

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+.n2p.model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US09774490/runatc_03082003_102929_4564/app_query.fasta.1.2887
-DB=Published Applications_AA -QMT=faetan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USPR=US09774490 @CGN 1.1.85 @runatc_03082003_102929_4564
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

71	86.1	771	15	US-10-097-340-284	Sequence 284, App
72	86.1	771	15	US-10-262-538-10	Sequence 10, App1
73	86.1	771	15	US-09-864-761-47112	Sequence 47112, A
74	5.9	53	9	US-09-864-761-44075	Sequence 44075, A
75	4.6	41	9	US-09-864-761-44552	Sequence 44552, A
76	3.7	33	9	US-09-864-761-35708	Sequence 35708, A
77	1.9	779	9	US-09-731-179-2	Sequence 2, App11
78	1.9	782	9	US-09-813-290-4	Sequence 4, App11
79	1.9	875	9	US-09-813-290-2	Sequence 12, App11
80	1.7	749	15	US-10-262-538-12	Sequence 214, App
81	1.7	751	14	US-10-114-833-214	Sequence 172, App
82	1.7	751	15	US-10-060-036-172	Sequence 359, App
83	1.7	751	15	US-10-205-823-359	Sequence 14, App1
84	1.7	751	15	US-10-262-538-14	Sequence 310, App
85	1.6	777	11	US-09-946-374-110	Sequence 348, App
86	1.6	777	14	US-10-052-586-348	Sequence 348, App
87	1.6	777	15	US-10-174-580-348	Sequence 348, App
88	1.6	777	15	US-10-176-758-348	Sequence 348, App
89	1.6	777	15	US-10-175-737-348	Sequence 348, App
90	1.6	777	15	US-10-173-706-348	Sequence 348, App
91	1.6	777	15	US-10-175-738-348	Sequence 348, App
92	1.6	777	15	US-10-175-738-348	Sequence 348, App
93	1.6	777	15	US-10-175-738-348	Sequence 348, App
94	1.6	777	15	US-10-175-738-348	Sequence 348, App
95	1.6	777	15	US-10-175-738-348	Sequence 348, App
96	1.6	777	15	US-10-175-738-348	Sequence 348, App
97	1.6	777	15	US-10-175-738-348	Sequence 348, App
98	1.6	777	15	US-10-175-738-348	Sequence 348, App
99	1.6	777	15	US-10-175-738-348	Sequence 348, App
100	1.6	777	15	US-10-175-738-348	Sequence 348, App

```

74 14 1.6 777 15 US-10-184-623-348 Sequence 348, App
75 14 1.6 777 15 US-10-184-635-348 Sequence 348, App
76 14 1.6 777 15 US-10-184-637-348 Sequence 348, App
77 14 1.6 777 15 US-10-184-646-348 Sequence 348, App
78 14 1.6 777 15 US-10-184-647-348 Sequence 348, App
79 14 1.6 777 15 US-10-184-652-348 Sequence 348, App
80 14 1.6 777 15 US-10-187-594-348 Sequence 348, App
81 14 1.6 777 15 US-10-187-596-348 Sequence 348, App
82 14 1.6 777 15 US-10-187-745-348 Sequence 348, App
83 14 1.6 777 15 US-10-187-885-348 Sequence 348, App
84 14 1.6 777 15 US-10-187-886-348 Sequence 348, App
85 14 1.6 777 15 US-10-199-464-348 Sequence 348, App
86 14 1.6 777 15 US-10-196-756-348 Sequence 348, App
87 14 1.6 777 15 US-10-176-751-348 Sequence 348, App
88 14 1.6 777 15 US-10-176-760-348 Sequence 348, App
89 14 1.6 777 15 US-10-176-990-348 Sequence 348, App
90 14 1.6 777 15 US-10-180-541-348 Sequence 348, App
91 14 1.6 777 15 US-10-180-542-348 Sequence 348, App
92 14 1.6 777 15 US-10-180-548-348 Sequence 348, App
93 14 1.6 777 15 US-10-180-551-348 Sequence 348, App
94 14 1.6 777 15 US-10-180-998-348 Sequence 348, App
95 14 1.6 777 15 US-10-180-999-348 Sequence 348, App
96 14 1.6 777 15 US-10-183-013-348 Sequence 348, App
97 14 1.6 777 15 US-10-184-612-348 Sequence 348, App
98 14 1.6 777 15 US-10-184-616-348 Sequence 348, App
99 14 1.6 777 15 US-10-184-617-348 Sequence 348, App
100 14 1.6 777 15 US-10-184-622-348 Sequence 348, App

```

ALIGNMENTS

RESULT 1

```

US-10-097-340-284
Sequence 284, Application US/10097340
Publication No. US20030087250A1

```

GENERAL INFORMATION:

```

APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VERIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

Alignment Scores:
Pred. No.: 0
Score: 771.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 86.15%
DB: 15
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGGCTGTTAACTGATGATGCTGCTCTTTCTGGGAGATTAATTACAGCAAGACA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuThrAlaArgAla 20

QY 260 AACTATCGAATGGGAGAGAACAAATGGCCAGGCTGAATTTCTTACAAAGAAATGTTG 319
DB 21 AsnTyrgIAsnGlyLVAsnAsnValProArgLeuLVLeuSerTyrgIAsnGlyLeu 40

QY 320 GAATCCAAACAATGTGATCATCTTTCATGAGCTTGCCCAAGCTCCAGTTATCATCCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrgIAsnThrPhe 60

QY 380 CTTTGGATGAGCAACGAGTAGGCTGTATGTTGAGCAACAAAGATCAATATTTTCATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTyrgValGlyAlaLVAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATATCAAGATTTTCAAAAGATTTGTCGCCAGTATCTTACACCAAGAGA 499
DB 81 AspLeuValAsnIleAspAspPheGlnIleValIleTrpProValSerTyrgIAsnArg 100

QY 500 GATGAATGCAATGGGCTGGAAGAAACATCCGTAAGAATGCTAATTTTCATCAAGATA 559
DB 101 AspGluCysGlyTrpAlaGlyIAspIleLeuGlyCysGlyAlaAsnPheIleVal 120

QY 560 CTTAAGCATATATATCAATCACTGATGACCTGTGAAACGGGCTTTTCATCAAT 619
DB 121 LeuValAlaTyrgAsnGlnThrHisLeuTyrgAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCTCATTTGAATATGAGCATATCATCTGAGCAATATTTTAAGTGGAGAACTCA 679
DB 141 CysTrpTyrgIleGlnIleGlyHisIleProGluAspAsnIlePheLVLeuLVAsnSer 160

QY 680 CATTTTGAAGACGGCGCTGGAGAGATCCATATGACCTTAAGCTGTCAGACATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyIAspSerProTyrgAspProLVLeuThrAlaSerIleu 180

QY 740 TTAATGATGAGAAATTAATTAATCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrgSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

QY 800 ATCTCCGAACCTCTGGGACACACACCAATCAGAGACAGACATGATTCAGAGTGG 859
DB 201 IlePheArgThrLeuGlyHisIleHisProIleArgGlnGlnHisAspSerTyrgTrp 220

QY 860 CTCATGATCCAAAGTTATGATGATGATCCCACTCATCTGAGAGATGACAAATCTGAAGAT 919
DB 221 LeuAsnAspProLVLeuPheIleSerHisIleLeuIleSerGluSerAspAsnProIleu 240

QY 920 GACAAAGTATCTTTTCTTCGTTGAAGAAATGCAATATGAGAAACACTCTGGAAGCT 979
DB 241 AspLVsValTyrgPhePhePheArgGluAsnAlaIleAspGlyGlnHisSerGlyLVsAla 260

QY 980 ACTCAGCTGAAGATGTCAGATATGACAAAGATGCTTTGAGGGCACAAGAGTTGCTGG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLVsAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAATGAGCAACAATCTCAAGCTGCTGATTTGCTCAGTCCAGGTCCAATGAC 1099

```


QY 500 GATGAATGCAAGTGGGCTGGAAAAAGATCTGTGAAGATGTGCTAATTTTCATCAAGTA 559
| | | | |
Db 101 AspGluCybelystPrpIaGlybAspIleLeuLysGluCybAlaAspHeiLeuYbVal 120
QY 560 CTTAAGCATATATATCAGATCTGCTGTACGCTGTGGAAAGGGGGCTTTTCATCCAAAT 619
| | | | |
Db 121 LeuLysAlaTyraAsnGlnThrAlaLeuYrAlaCybGlyThrGlyAlaPheHeiSproIle 140
QY 620 TGCACCTACATTAATAATGGACATCATCTGAGAGACAATTTTAACTGAGAGACTCA 679
| | | | |
Db 141 CysThrTyrlleGluIleGlyAlaSH1SproGluAspAsnIlePheYbSleuGluAsnSer 160
QY 680 CATTTTGAACACGGCGGTGGAGAGATCATATGACCTTAAGCTGTGACAGCATCTT 739
| | | | |
Db 161 HisPheGluAsnGlybArgGlyYbSerProTyraAspProYbSleuThrAlaSerLeu 180
QY 740 TTAAATGATGAGAAATTAATCTGTGAATGCACTGAGTATTATAGGGCGAGACTTTGCT 799
| | | | |
Db 181 LeuIleAspGlyGlyLeuTyraSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGCAACTCTTGGGACCAACCCCAATGAGACAGAGAGATGATTCAGAGTG 859
| | | | |
Db 201 IlePheArgThrLeuGlybSH1SH1SproIleArgThrGlyGlnIbAspSerArgTyP 220
QY 860 CTCATGATCCAAAGTTTCAATTAAGTCCCACTCATCTCAGAGAGTGAACAATCCTGAAGAT 919
| | | | |
Db 221 LeuAsnAspProYbSleuHeiLeSerAlaHeiLeuIleSerGluSerbAspAsnProGluAsp 240
QY 920 GACAAAGTATACTTTTCTTCGCTGAATGCAATGATGAGAGACAATCTGTGAAAAAGCT 979
| | | | |
Db 241 AspYbValTyraPhePhePheArgGluAsnAlaIleAspGlyGlnIbSerserGlybAla 260
QY 980 ACTCAGCTAGATATAGTATGCAAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1039
| | | | |
Db 261 ThrH1SalArgIleGlyGlnIleCybYbAsnAspPheGlyGlybArgSerLeuVal 280
QY 1040 AATAATGAGACAATCTCTCAAGCTCGTGTGATTTGCTCAGTGCAGAGTCCCAATGAG 1099
| | | | |
Db 281 AsnLysTrpThrThrPheLeuYbAlaGlyLeuIleCybSerValProGlyProAsnGly 300
QY 1100 ATTGACATCTCATTTTGAATGAACTGCAAGATGATTCCTTAATGAACTTTAAAGATCTTAA 1159
| | | | |
Db 301 IleAspThrHisPheAspGlyLeuGlnAspValPheLeuMetAsnPheYbAspProYbS 320
QY 1160 AATCCAGTTGATATGAGAGTGTAGACTTCAGCTTCAGATTAATTTCAAGGATCCAGCG 1219
| | | | |
Db 321 AsnProValValTyrlValPheThrThrSerSerAsnIlePheYbSglYserAlaVal 340
QY 1220 TGTATGTATAGCATGATGATGAGAGAGGTTGCTTGTGTCATATGCCACAGAGAT 1279
| | | | |
Db 341 CysMetTyraSerMetSerAspValArgValPheLeuGlyProTyraAlaHisArgAsp 360
QY 1280 GGAACCACTATCATAGGCTGCTTATCAAGAGAGATCCCTATCCACGGCCAGAGACT 1339
| | | | |
Db 361 GlyProAsnTyrlGlnTrpValProTyrlGlnGlyArgValProTyraProArgProGlyThr 380
QY 1340 TGTCCAGAGAAACATTTGTGTGTTTGAATCTCAACAAGAGACCTTCTGATGATGTTAA 1399
| | | | |
Db 381 CysProSerTyraPheGlyGlyPheAspSerThrYbAspLeuProAspValIle 400
QY 1400 ACCTTTGAAGAAGTATCATCAGCATGATCAATCCAGTGTTCCTATGAACATCGCCCA 1459
| | | | |
Db 401 ThrPheAlaArgSerHisProAlaMetTyraAsnProValPheProMetAsnAsnAspPro 420
QY 1460 ATAGTATCAAAACGAGATGTAATTAATCAATTTTACAAATTTGCTAGACCGAGTGAAT 1519
| | | | |
Db 421 IleValIleTyraThrAspValAsnTyrlGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GAGAGAGATGAGACAGATGATGTTATGCTGAGAGAGAGATGTTGGACCGTCTTAA 1579
| | | | |
Db 441 AlaGluAspGlyGlnTyraAspValMetPheIleGlyThrAspValGlyThrValLeuYbS 460

QY 1580 GTAGTTTCATTTCTTAAGAGAGACTTGATGATTTAGAAAGGTTCTGTGGAGAAAAATG 1639
| | | | |
Db 461 ValIleSerIleProYbGlnThrTrpTyraPheLeuGlnGluValIleLeuGlnGluMet 480
QY 1640 ACAGTTTTTGGGAAACCGACTGCTATTTTACAGAAATGAGAGCTTTCCATTAAGAGACAACA 1699
| | | | |
Db 481 ThrValPheArgGlybProThrAlaIleSerAlaMetGluLeuSerThrYbSglGlnGln 500
QY 1700 CTATATATTTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGATGATATTAC 1759
| | | | |
Db 501 LeuTyrlleGlySerThrAlaGlyAlaAlaGlnLeuProLeuHisArgCybAspIleTyR 520
QY 1760 GGAAGAACGCTGTGCTAGTGTTCCTCCCGCAGACCTTACTGTGCTTGGAGTGTCT 1819
| | | | |
Db 521 GlyYbAlaCybAlaGluCybSleuAlaArgAspProTyraValAlaTrpAspGlySer 540
QY 1820 GATGTTCTCGTATTTTCCCTGCAAGAGAGACCAACAAGCAACAAGATTAAGAAT 1879
| | | | |
Db 541 AlaCysSerArgTyraPheProThrAlaYbArgArgTrpArgGlnAspIleArgAsn 560
QY 1880 GGAAGACCCAGTCACTCATGTTCAAGCTTACACCATGATTAATCAACATGGCCAGAGCCCT 1939
| | | | |
Db 561 GlyAspProLeuThrHisCybSerAspLeuHisSH1SAspAsnHisSH1GlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGTGTAGAGAAATAGTACCAATTTTGAATGCAATGCCAGAG 1999
| | | | |
Db 581 GlnGluArgIleIleTyrlGlyValGluAsnSerSerThrPheLeuGluCybSerProYbS 600
QY 2000 TGGCAGAGAGCGCTGTCTATTTGGCAATCCAGAGAGCGAAATGAAGAGCGAAAGAAAGAG 2059
| | | | |
Db 601 SerGlnArgAlaLeuValTyrlTrpGlnPheGlnArgAlaGlyAsnGlnGluArgGlyGln 620
QY 2060 ATCAGAGTGAATGATCATATCATCATGAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
| | | | |
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGATTCAGGCAATTAACCTCTCCATGCGGTGAGAAACATGGGTTATCAAACTGT 2179
| | | | |
Db 641 GlnYbAspSerGlybAsnTyraLeuYbSH1SH1AlaValGlnIbSgLYPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCTGGAAGTCAATGACACAGAGCATTTTGAAGAACTTTCTATTAAGAT 2239
| | | | |
Db 661 LeuYbValThrLeuGlnValIleAspThrGlnHisLeuGlnGluLeuLeuHisYbAsp 680
QY 2240 GATGATGAGAGATGCTTCAAGCCAAAGAAATGTCATATGATGATGACCTAGCCAGAG 2299
| | | | |
Db 681 AspAspGlybAspGlySerTyraSerTyraSerLeuMetSerbAsnSerMetThrProSerGlnYbS 700
QY 2300 GTCTGTACAGAGACTTATGATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGAG 2359
| | | | |
Db 701 ValTrpTyraArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAATTTTGAAGAAAGGACCGAAACAAGTCCGCAAAAGGCCAGAGCATACC 2419
| | | | |
Db 721 PheCysGlnGlnValTrpTyraArgAspArgYbGlnArgGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGAGAGACTTACAAGAAATTAAGAAAGTAAAGAAACGAGAGG 2479
| | | | |
Db 741 ProGlyAsnSerAsnLysTrpTyraSH1SleuGlnGluAsnLysYbSgLYArgAsnArgArg 760
QY 2480 ACCCAGAAATTTGAGAGGAGCAACCCAGAGAGTGC 2512
| | | | |
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3
US-09-864-761-47112
; Sequence 47112, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

```

      TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
      TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
      FILE REFERENCE: Aeonica-X-1
      CURRENT FILING DATE: 2001-05-23
      PRIOR APPLICATION NUMBER: US 09/180,312
      PRIOR FILING DATE: 2000-02-04
      PRIOR APPLICATION NUMBER: US 60/207,456
      PRIOR FILING DATE: 2000-05-26
      PRIOR APPLICATION NUMBER: US 09/632,366
      PRIOR FILING DATE: 2000-08-03
      PRIOR APPLICATION NUMBER: GB 24263.6
      PRIOR FILING DATE: 2000-10-04
      PRIOR APPLICATION NUMBER: US 60/236,359
      PRIOR FILING DATE: 2000-09-27
      PRIOR APPLICATION NUMBER: PCT/US01/00666
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00667
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00664
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00669
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00665
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00668
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00663
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00662
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00661
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: PCT/US01/00670
      PRIOR FILING DATE: 2001-01-30
      PRIOR APPLICATION NUMBER: US 60/234,687
      PRIOR FILING DATE: 2000-09-21
      PRIOR APPLICATION NUMBER: US 09/608,408
      PRIOR FILING DATE: 2000-06-30
      PRIOR APPLICATION NUMBER: US 09/774,203
      PRIOR FILING DATE: 2001-01-29
      NUMBER OF SEQ ID NOS: 49117
      SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
      SEQ ID NO 47112
      LENGTH: 151
      TYPE: PRT
      ORGANISM: Homo sapiens
      FEATURE:
      OTHER INFORMATION: MAP TO AC006322.2
      OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
      OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
      OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU8 2.00e-88
      US-09-864-761-47112

Alignment Scores:
Pred. No.: 1.39e-140 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-47112 (1-151)

QY 2060 ATCAAGTGTGATGATATATCATCAGAGACATCAAGGCGCTTCGTAAGTACTCAA 2119
DB 1 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
   11leaGvaAAspAspHisIleIleArgThAspGlnGlyLeuLeuLeuAAsGSeLeuGln 20

QY 2120 CAGAAGATTCCAGCATTAACCTCTGCATCCGCGTGGAGATGGATTATACAAACTTT 2179
DB 21 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
   21GlnYAspSerGlySerGlyLeuCYHisAlaValAluIuHisGlyPheIleGlnThIleu 40

2180 CTTAAAGCTAACCTCGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATTAAGAT 2239

```

Db 41 LeuysValThrLeuGlnValIleAspThrGlnHisLeuGlnIleuLeuHisIleValAsp 60

Qy 2240 GATGATGAGAAAGGCTCTTAAGACCAAGAAATGTCCAAATAGCATGACACCTAGCCAGAG 2239

Db 61 AspaBgIlyAspGlySerIyThrIyGlnMetSerAsnSerMetIhrProSerGlnIly 80

Qy 2300 GTCTGGTACAGAGACTTTCATGACAGCTCATCAACCCCAATCTCAACACGATGATGAG 2359

Db 81 ValTrpIyTrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 100

Qy 2360 TTCTGTGAACAAGTTTGGAAAAAGGAGCGAAAAACAAGTCGGCAAGGCCAGGACATACC 2419

Db 101 PheCysGlnGlnValTrpIyAsnArgAspArgIlyGlnArgGlnIhrProGlyHisThr 120

Qy 2420 CCAGGAACAGTACCAAAATGAGACCTTACAGAAATAAGAAAGTAGAAAAAGAGAG 2479

Db 121 ProGlyAsnSerAsnIySerTrpIyShIleuGlnGlnAsnIlySerGlyArgAsnArgArg 140

Qy 2480 ACCACGAAATTTGAAGAGGCGACCCAGAGATGTC 2512

Db 141 ThrHisGlnPheGlnArgAlaProArgSerVal 151

RESULT 4

US-09-864-761-44075

Sequence 44075, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

```

NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 44075
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU8 5.00e-25
US-09-864-761-44075

Alignment Scores:
Pred. No.: 2.73e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-44075 (1-57)

QY 311 GAAATGTTGGAATCCAAACATGTGATCATTTCATGAGCTTGCGCAACAGCTCCAGTTAT 370
Db 5 G1umellenglUsermbnenvalllethPheasnolyleuAlaanserSertyr 24

QY 371 CATACCTTCCTTTTGGATGAGAACGAGTAGCGCTGATGTTGGAGCAAAAGATCACATA 430
Db 25 H1strhPheuleuaspGingulnArgserArgleutyryValGlyAlalyAspHisle 44

QY 431 TTTTCATTCGACCTGCTTAATATCAAGGATTTTCAAAG 469
Db 45 PheSerPheaspLeuvalAlaHllelyAspPheGlnlys 57

RESULT 5
US-09-864-761-44552
Sequence 44552, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30

```

```

PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 44552
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
OTHER INFORMATION: EST HUMAN HIT: BF700780.1, VALUE 1.00e-17
OTHER INFORMATION: SWISSPROT HIT: Q14563, VALUE 1.00e-18
US-09-864-761-44552

Alignment Scores:
Pred. No.: 2,29e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)

QY 1937 CCGAAGAGAATCATTCATGTGTGAGAAATAGACATTTTGGAATGCAGTCCG 1996
Dn 1 ProglunlurlylelletryglvalGlubenseSerthPhelengluyssePro 20
QY 1997 AAGTCGACAGAGAGCGCTGCTATTGGCAATTCACAGGCCGAATGAAGCAGAAAAGA 2056
Dn 21 LysercGlurghaleuValTyrrtgcInphedlnArgrAngluclurglyslu 40
QY 2057 GAG 2059
Dn 41 Glu 41

RESULT 6
US-09-864-761-39708
Sequence 39708 Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenhong
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03

```

```
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 39708
/ LENGTH: 33
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 4.00e-15
US-09-864-761-39708

Alignment Scores:
Pred. No.: 2,03e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-39708 (1-33)

QY 533 AAGAGTGTCTTAATTTTCATCAGAGTACTTAAGCATATATACAGACTTGTAGCC 592
Db 1 LysGUCysAlaAsnPhelleySvalleySAlaTyAsnGlnThrHisLeuTyAla 20
QY 593 TGTGACGGGGGCTTTTCATTCATTCATTTGACCTTACATT 631
Db 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 7
US-09-731-179-2
/ Sequence 2, Application US/09731179
/ Patent No. US20010049432A1
/ GENERAL INFORMATION:
/ APPLICANT: Holloway, James L..
/ APPLICANT: Foley, Kevin P.
```

```
/ TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
/ FILE REFERENCE: 99-90
/ CURRENT APPLICATION NUMBER: US/09/731,179
/ CURRENT FILING DATE: 2000-12-06
/ PRIOR APPLICATION NUMBER: US 60/169,238
/ PRIOR FILING DATE: 1999-12-06
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 779
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-731-179-2

Alignment Scores:
Pred. No.: 1.11e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-731-179-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGAGATGT 1816
Db 525 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIlePaspGly 541

RESULT 8
US-09-813-290-4
/ Sequence 4, Application US/09813290
/ Patent No. US20020042504A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Wilganowski, Nathaniel L.
/ APPLICANT: Turner, C. Alexander Jr.
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Wang, Xiaoming
/ APPLICANT: Donoho, Gregory
/ APPLICANT: Scottille, John
/ TITLE OF INVENTION: NO. US20020042504A1e1 Human Secreted Proteins and Polynucleotides
/ FILE REFERENCE: LEX-0151-USA
/ CURRENT APPLICATION NUMBER: US/09/813,290
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: US 60/190,638
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: US 60/191,188
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: US 60/193,639
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-09-813-290-4

Alignment Scores:
Pred. No.: 1.11e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-4 (1-782)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGAGATGT 1816
Db 528 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIlePaspGly 544

RESULT 9
```

```
US-09-813-290-2
; Sequence 2, Application US/09813290
; Patent No. US20020042504A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Hilbun, Erin
; APPLICANT: Wang, Xiaoming
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. US20020042504A1 Human Secreted Proteins and Polynucleotides
; FILE REFERENCE: Lex-0151-USA
; CURRENT APPLICATION NUMBER: US/09/813,290
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,638
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: US 60/191,188
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/193,639
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 875
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-813-290-2

Alignment Scores:
Pred. No.: 1,1e-07 Length: 875
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-2 (1-875)
QY 1766 GGGTGTGGAGTGGTGGCCGCGGAGCCCTTACTGTCGTTGGAGTGT 1816
DB 621 AAlAcysAlaGluCysCysLeuAlaArgApProtyrCyAlaIrrpAspGly 637

RESULT 10
US-10-262-538-12
; Sequence 12, Application US/10262538
; Publication No. US20030113324A1
; GENERAL INFORMATION:
; APPLICANT: Altalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-12

Alignment Scores:
Pred. No.: 1,08e-05 Length: 749
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-12 (1-749)
QY 1028 AGAGCTGTGTAATGAATGACAAACATTCTCAAGCTGCTG 1072
DB 274 ArgSerLeuValAsnLyseTrrThrPheLeuLysAlaArgLeu 288

RESULT 11
US-10-114-893-214
; Sequence 214, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaValle, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-214

Alignment Scores:
Pred. No.: 1,08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 14 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-114-893-214 (1-751)
QY 1028 AGAGCTGTGTAATGAATGACAAACATTCTCAAGCTGCTG 1072
DB 274 ArgSerLeuValAsnLyseTrrThrPheLeuLysAlaArgLeu 288

RESULT 12
US-10-060-036-172
; Sequence 172, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-172

Alignment Scores:
Pred. No.: 1,08e-05 Length: 751
```

Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-060-036-172 (1-751)

QY 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
|||||
DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 13

US-10-205-823-359
Sequence 359, Application US/10205823
Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbatcheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITILE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25

PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 359
LENGTH: 751
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-359

Alignment Scores:

Pred. No.: 1.08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-205-823-359 (1-751)

QY 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
|||||
DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 14

US-10-262-538-14
Sequence 14, Application US/10262538
Publication No. US20030113324A1

GENERAL INFORMATION:
APPLICANT: Altio et al
TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS

FILE REFERENCE: 28967/37564
CURRENT APPLICATION NUMBER: US/10/262,538
CURRENT FILING DATE: 2002-09-30
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 751
TYPE: PRT
ORGANISM: Homo sapiens
US-10-262-538-14

Alignment Scores:

Pred. No.: 1.08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-14 (1-751)

QY 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
|||||
DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 15

US-09-946-374-310
Sequence 310, Application US/09946374
Publication No. US20030073129A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids
FILE REFERENCE: P2830P1C1
CURRENT APPLICATION NUMBER: US/09/946,374
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596

[illegible]

;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807
Alignment Scores:
Pred. No.: 0 000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 11 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-946-374-310 (1-777)
Qy 1040 AATAATGACAACTTCTCAAGCTCGTCAATTGCTCA 1081
Db 298 AenlySTPThrPhelenuySalAargLeuIleCySer 311
RESULT 16
US-10-052-586-348
; Sequence 348, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052.586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065111
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120

;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/066466
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/066772
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/069335
;; PRIOR FILING DATE: 1997-12-11
;; PRIOR APPLICATION NUMBER: 60/069425
;; PRIOR FILING DATE: 1997-12-12
;; PRIOR APPLICATION NUMBER: 60/069870
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/068017
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/07450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15

;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/086332
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088336
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538

;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 14 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-052-586-348 (1-777)

Qy 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlySTpThrThrPheuleylsalargleuileCysser 311

RESULT 17
US-10-174-590-348

;; Sequence 348, Application US/10174590
;; Publication No. US20030008352A1

;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.

;; APPLICANT: Chen, Jian

;; APPLICANT: Desnoyers, Luc

;; APPLICANT: Goddard, Audrey

;; APPLICANT: Godowski, Paul J.

;; APPLICANT: Gurney, Austin L.

;; APPLICANT: Pan, James

;; APPLICANT: Smith, Victoria

;; APPLICANT: Matanabe, Colin K.

;; APPLICANT: Wood, William I.

;; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

;; FILE REFERENCE: P3430R1C42

;; CURRENT APPLICATION NUMBER: US/10/174,590

;; CURRENT FILING DATE: 2002-06-18

;; Prior application removed - See file wrapper or Palm

;; NUMBER OF SEQ ID NOS: 612

;; SEQ ID NO 348

;; LENGTH: 777

;; TYPE: PRT

;; ORGANISM: Homo Sapien

US-10-174-590-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-590-348 (1-777)

Qy 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlySTpThrThrPheuleylsalargleuileCysser 311

RESULT 18
US-10-176-758-348

;; Sequence 348, Application US/10176758
;; Publication No. US20030008353A1

;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.

;; APPLICANT: Chen, Jian

;; APPLICANT: Desnoyers, Luc

;; APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-758-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-758-348 (1-777)

Qy 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpnrThrPheleuylsAlaArgleuileCysSer 311

RESULT 19
US-10-175-737-348
Sequence 348, Application US/10175737
Publication No. US20030013153A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C50
CURRENT APPLICATION NUMBER: US/10/175,737
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-737-348 (1-777)

Qy 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpnrThrPheleuylsAlaArgleuileCysSer 311

RESULT 20
US-10-173-706-348
Sequence 348, Application US/10173706
Publication No. US20030022293A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-706-348 (1-777)

Qy 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpnrThrPheleuylsAlaArgleuileCysSer 311

RESULT 21
US-10-175-738-348
Sequence 348, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT

```

; ORGANISM: Homo Sapien
US-10-175-738-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-738-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AsnLysTrpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 22
US-10-175-752-348
; Sequence 348, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-752-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AsnLysTrpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 23
US-10-176-482-348
; Sequence 348, Application US/10176482
; Publication No. US20030022286A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
```

```

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-482-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AsnLysTrpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 24
US-10-176-757-348
; Sequence 348, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-757-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AsnLysTrpThrPheLeuYsAlaArgLeuIleCySer 311
```

```
RESULT 25
US-10-176-913-348
; Sequence 348, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-913-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGTCGATTGCTCA 1081
DB 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 26
US-10-180-552-348
; Sequence 348, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-348

Alignment Scores:
```

```
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-552-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGTCGATTGCTCA 1081
DB 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 27
US-10-180-557-348
; Sequence 348, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-557-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGTCGATTGCTCA 1081
DB 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 28
US-10-173-700-348
; Sequence 348, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-700-348 (1-777)

QY 1040 AATTAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AaNLystrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 29
US-10-174-572-348
; Sequence 348, Application US/10174572
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-572-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-572-348 (1-777)

QY 1040 AATTAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AaNLystrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 30
US-10-174-579-348
; Sequence 348, Application US/10174579
; Publication No. US20030027264A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-579-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-579-348 (1-777)

QY 1040 AATTAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AaNLystrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 31
US-10-174-582-348
; Sequence 348, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-582-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```



```
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-588-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
298 AenlystirpThrPheleuysAlaArgLeuIleCySer 311

RESULT 32
US-10-174-588-348
; Sequence 348, Application US/10174588
; Publication No. US200300272661
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C28
; CURRENT APPLICATION NUMBER: US/10/174,588
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-588-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-588-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
298 AenlystirpThrPheleuysAlaArgLeuIleCySer 311

RESULT 33
US-10-175-739-348
; Sequence 348, Application US/10175739
; Publication No. US20030027267A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C46
; CURRENT APPLICATION NUMBER: US/10/175,739
; CURRENT FILING DATE: 2002-06-19
```

```
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-739-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-175-739-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
298 AenlystirpThrPheleuysAlaArgLeuIleCySer 311

RESULT 34
US-10-175-740-348
; Sequence 348, Application US/10175740
; Publication No. US20030027268A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C61
; CURRENT APPLICATION NUMBER: US/10/175,740
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-740-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-175-740-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
298 AenlystirpThrPheleuysAlaArgLeuIleCySer 311

RESULT 35
US-10-175-743-348
; Sequence 348, Application US/10175743
; Publication No. US20030027269A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
```

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C52
CURRENT FILING DATE: 2002-06-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/08582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-743-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 238 AenlystPthThrPheleuylsAlaArgleuileCysSer 311

RESULT 36
US-10-176-488-348
Sequence 348, Application US/10176488
Publication No. US20030027271A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C119
CURRENT APPLICATION NUMBER: US/10/176,488
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-488-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-488-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 238 AenlystPthThrPheleuylsAlaArgleuileCysSer 311

RESULT 37
US-10-176-492-348
Sequence 348, Application US/10176492
Publication No. US20030027272A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C107
CURRENT APPLICATION NUMBER: US/10/176,492
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-492-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-492-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 38

US-10-176-747-348
; Sequence 348, Application US/10176747
; Publication No. US20030027278A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C92
; CURRENT APPLICATION NUMBER: US/10/176,747
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-747-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-747-348 (1-777)

QY 1040 AATAAATGACAACATTCCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 39

US-10-176-750-348
; Sequence 348, Application US/10176750
; Publication No. US20030027274A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C103
; CURRENT APPLICATION NUMBER: US/10/176,750
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-176-750-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-750-348 (1-777)

QY 1040 AATAAATGACAACATTCCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 40

US-10-176-985-348
; Sequence 348, Application US/10176985
; Publication No. US20030027277A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C99
; CURRENT APPLICATION NUMBER: US/10/176,985
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-985-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-985-348 (1-777)

QY 1040 AATAAATGACAACATTCCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 41

US-10-176-987-348
; Sequence 348, Application US/10176987
; Publication No. US20030027278A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.

APPLICANT: Wood,William I.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P343ORIC93
CURRENT APPLICATION NUMBER: US/10/176,987
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-987-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-987-348 (1-777)

QY 1040 AATAATGACAAACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiprthrPheleuylsAlaArgleuileCysSer 311

RESULT 42
US-10-176-992-348
Sequence 348, Application US/10176992
Publication No. US20030027279A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P343ORIC100
CURRENT APPLICATION NUMBER: US/10/176,992
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-992-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-992-348 (1-777)

QY 1040 AATAATGACAAACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiprthrPheleuylsAlaArgleuileCysSer 311
RESULT 43

US-10-176-993-348
Sequence 348, Application US/10176993
Publication No. US20030027280A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P343ORIC89
CURRENT APPLICATION NUMBER: US/10/176,993
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-993-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-993-348 (1-777)

QY 1040 AATAATGACAAACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiprthrPheleuylsAlaArgleuileCysSer 311

RESULT 44
US-10-184-658-348
Sequence 348, Application US/10184658
Publication No. US20030027281A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P343ORIC228
CURRENT APPLICATION NUMBER: US/10/184,658
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-658-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-184-658-348 (1-777)

QY 1040 AATAATGACACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheuleuysAlaArgLeuileCysser 311

RESULT 45

US-10-176-991-348
; Sequence 348, Application US/10176991
; Publication No. US20030027324A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C122
; CURRENT APPLICATION NUMBER: US/10/176,991
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-991-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-991-348 (1-777)

QY 1040 AATAATGACACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheuleuysAlaArgLeuileCysser 311

RESULT 46

US-10-173-695-348
; Sequence 348, Application US/10173695
; Publication No. US20030032101A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C122
; CURRENT APPLICATION NUMBER: US/10/173,695
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-695-348

FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,695
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-695-348

QY 1040 AATAATGACACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheuleuysAlaArgLeuileCysser 311

RESULT 47

US-10-173-697-348
; Sequence 348, Application US/10173697
; Publication No. US20030032102A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,697
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-697-348

QY 1040 AATAATGACACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheuleuysAlaArgLeuileCysser 311

RESULT 48

US-10-173-705-348
; Sequence 348, Application US/10173705
; Publication No. US20030032103A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,705
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-705-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-697-348 (1-777)

QY 1040 AATAATGACACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheuleuysAlaArgLeuileCysser 311

RESULT 48

US-10-173-705-348
; Sequence 348, Application US/10173705
; Publication No. US20030032103A1
; GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C18
; CURRENT APPLICATION NUMBER: US/10/173,705
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-705-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-705-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 Aenlystptpnrthrpheulysalargleuilecysser 311

RESULT 49
US-10-174-576-348
; Sequence 348, Application US/10174576
; Publication No. US20030032104A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C23
; CURRENT APPLICATION NUMBER: US/10/174,576
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-576-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
```

```

DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-576-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 Aenlystptpnrthrpheulysalargleuilecysser 311

RESULT 50
US-10-174-585-348
; Sequence 348, Application US/10174585
; Publication No. US20030032105A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C17
; CURRENT APPLICATION NUMBER: US/10/174,585
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-585-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-585-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 Aenlystptpnrthrpheulysalargleuilecysser 311

RESULT 51
US-10-174-586-348
; Sequence 348, Application US/10174586
; Publication No. US20030032106A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C24
; CURRENT APPLICATION NUMBER: US/10/174,586
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
```



```
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-174-586-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-586-348 (1-777)

QY 1040 AATAAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
| | | | | | | | | | | | | | | | | | | | | |
Db 298 Asnlystrpthrphreuleysalargleuilecysser 311

RESULT 52
US-10-175-747-348
/ Sequence 348, Application US/10175747
/ Publication No. US20030032107A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Matanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C44
/ CURRENT APPLICATION NUMBER: US/10/175,747
/ CURRENT FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063564
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063734
/ PRIOR FILING DATE: 1997-10-29
/ PRIOR APPLICATION NUMBER: 60/063870
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/064103
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066120
```

```
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/066466
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/066772
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/069335
/ PRIOR FILING DATE: 1997-12-11
/ PRIOR APPLICATION NUMBER: 60/069425
/ PRIOR FILING DATE: 1997-12-12
/ PRIOR APPLICATION NUMBER: 60/069870
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/068017
/ PRIOR FILING DATE: 1997-12-18
/ PRIOR APPLICATION NUMBER: 60/077450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/082568
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082569
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082704
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/082797
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/083495
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083496
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/084366
/ PRIOR FILING DATE: 1998-05-05
/ PRIOR APPLICATION NUMBER: 60/084414
/ PRIOR FILING DATE: 1998-05-06
/ PRIOR APPLICATION NUMBER: 60/084639
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084640
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084643
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/085573
/ PRIOR FILING DATE: 1998-05-15
```

```

; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538

```

```

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653

```

```

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-175-747-348 (1-777)

```

QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystfptThrPheleuylsAlaArgleuileCysSer 311

```

RESULT 53

```

US-10-176-481-348
; Sequence 348, Application US/10176481
; Publication No. US20030032108A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C98
; CURRENT APPLICATION NUMBER: US/10/176,481
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-176-481-348

```

Alignment Scores:

```

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-176-481-348 (1-777)

```

QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystfptThrPheleuylsAlaArgleuileCysSer 311

```

RESULT 54

```

US-10-176-485-348
; Sequence 348, Application US/10176485
; Publication No. US20030032109A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

```

APPLICANT: Pan,James
APPLICANT: Smith,Victoria
APPLICANT: Watanabe,Colin K.
APPLICANT: Wood,William I.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C78
CURRENT APPLICATION NUMBER: US/10/176,485
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-485-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-485-348 (1-777)

QY 1040 AATAATGACAACATTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCySer 311

RESULT 55

US-10-176-487-348
Sequence 348, Application US/10176487
Publication No. US20030032110A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C74
CURRENT APPLICATION NUMBER: US/10/176,487
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-487-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-487-348 (1-777)

QY 1040 AATAATGACAACATTCTCAAGCTGCTGATTGCTCA 1081
|||||
|||||

Db 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCySer 311

RESULT 56

US-10-176-493-348
Sequence 348, Application US/10176493
Publication No. US20030032111A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C72
CURRENT APPLICATION NUMBER: US/10/176,493
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-493-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-493-348 (1-777)

QY 1040 AATAATGACAACATTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCySer 311

RESULT 57

US-10-176-756-348
Sequence 348, Application US/10176756
Publication No. US20030032112A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C109
CURRENT APPLICATION NUMBER: US/10/176,756
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-756-348

QY 1040 AATAATGACAACATTCTCAAGCTGCTGATTGCTCA 1081
|||||
|||||

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-756-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 58
US-10-176-911-348

/ Sequence 348, Application US/10176911
/ Publication No. US20030032113A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C75
/ CURRENT APPLICATION NUMBER: US/10/176,911
/ CURRENT FILING DATE: 2002-06-20
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-911-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-911-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 59
US-10-176-919-348

/ Sequence 348, Application US/10176919
/ Publication No. US20030032114A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C63
/ CURRENT APPLICATION NUMBER: US/10/176,919
/ CURRENT FILING DATE: 2002-06-20
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-919-348

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 60
US-10-176-925-348

/ Sequence 348, Application US/10176925
/ Publication No. US20030032115A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C94
/ CURRENT APPLICATION NUMBER: US/10/176,925
/ CURRENT FILING DATE: 2002-06-21
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-925-348

US-09-774-490-1 (1-2709) x US-10-176-919-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 61
US-10-176-978-348

/ Sequence 348, Application US/10176928
/ Publication No. US20030032116A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C94
/ CURRENT APPLICATION NUMBER: US/10/176,928
/ CURRENT FILING DATE: 2002-06-21
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-928-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-925-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 61
US-10-176-978-348

```
/ Sequence 348, Application US/10176978
/ Publication No. US20030032116A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C16
/ CURRENT FILING DATE: 2002-06-21
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-978-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-978-348 (1-777)
QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 62
US-10-179-510-348
/ Sequence 348, Application US/10179510
/ Publication No. US20030032117A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C18
/ CURRENT FILING DATE: 2002-06-24
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-179-510-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-179-510-348 (1-777)
QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 63
US-10-180-543-348
/ Sequence 348, Application US/10180543
/ Publication No. US20030032118A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C161
/ CURRENT FILING DATE: 2002-06-25
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-180-543-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-543-348 (1-777)
QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 64
US-10-180-544-348
/ Sequence 348, Application US/10180544
/ Publication No. US20030032119A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C150
```

```

; CURRENT APPLICATION NUMBER: US/10/180,544
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-544-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-544-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
Db 298 AsnysrtpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 65
US-10-180-546-348
; Sequence 348, Application US/10180546
; Publication No. US20030032120A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C156
; CURRENT APPLICATION NUMBER: US/10/180,546
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-546-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-546-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
Db 298 AsnysrtpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 66
US-10-180-547-348
; Sequence 348, Application US/10180547
; Publication No. US20030032121A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
```

```

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C157
; CURRENT APPLICATION NUMBER: US/10/180,547
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-547-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-547-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
Db 298 AsnysrtpThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 67
US-10-180-549-348
; Sequence 348, Application US/10180549
; Publication No. US20030032122A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C151
; CURRENT APPLICATION NUMBER: US/10/180,549
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-549-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-180-549-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuylsAlaArgLeuileCysSer 311

RESULT 68

US-10-180-555-348

/ Sequence 348, Application US/10180555

/ Publication No. US20030032123A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C163

/ CURRENT APPLICATION NUMBER: US/10/180,555

/ CURRENT FILING DATE: 2002-06-25

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 612

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-180-555-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-555-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuylsAlaArgLeuileCysSer 311

RESULT 69

US-10-180-559-348

/ Sequence 348, Application US/10180559

/ Publication No. US20030032124A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C159

/ CURRENT APPLICATION NUMBER: US/10/180,559

/ CURRENT FILING DATE: 2002-06-25

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-180-559-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-559-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuylsAlaArgLeuileCysSer 311

RESULT 70

US-10-181-000-348

/ Sequence 348, Application US/10181000

/ Publication No. US20030032125A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C177

/ CURRENT APPLICATION NUMBER: US/10/181,000

/ CURRENT FILING DATE: 2002-06-26

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 612

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-181-000-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-181-000-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuylsAlaArgLeuileCysSer 311

RESULT 71

US-10-183-010-348

/ Sequence 348, Application US/10183010

/ Publication No. US20030032126A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.


```
/ APPLICANT: Gurney,Austin L.
/ APPLICANT: Pan,James
/ APPLICANT: Smith,Victoria
/ APPLICANT: Watanabe,Colin K.
/ APPLICANT: Wood,William I.
/ APPLICANT: Zhang,Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C164
/ CURRENT APPLICATION NUMBER: US/10/183,010
/ CURRENT FILING DATE: 2002-06-26
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-183-010-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-183-010-348 (1-777)
QY 1040 AATTAATGACAACTTCTCTCAAGCTGCTGATTTGCTCA 1081
Db 238 AasnlystipThrPhreuleysalaxgleullecyser 311

RESULT 72
US-10-183-012-348
/ Sequence 348, Application US/10183012
/ Publication No. US20030032127A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C171
/ CURRENT APPLICATION NUMBER: US/10/183,012
/ CURRENT FILING DATE: 2002-06-26
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
```

```
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063564
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063734
/ PRIOR FILING DATE: 1997-10-29
/ PRIOR APPLICATION NUMBER: 60/063870
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/064103
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066120
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/066466
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/066772
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/069335
/ PRIOR FILING DATE: 1997-12-11
/ PRIOR APPLICATION NUMBER: 60/069425
/ PRIOR FILING DATE: 1997-12-12
/ PRIOR APPLICATION NUMBER: 60/069870
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/068017
/ PRIOR FILING DATE: 1997-12-18
/ PRIOR APPLICATION NUMBER: 60/07450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/082568
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082569
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082704
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/082797
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/083495
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083496
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
```

```

PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088728
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863

```

```

PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

```

Alignment Scores:

```

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-183-012-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 73

```

US-10-184-614-348
Sequence 348, Application US/10184614
Publication No. US20030032128A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C184
CURRENT APPLICATION NUMBER: US/10/184,614
CURRENT FILING DATE: 2225-06-27
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-614-348

```

Alignment Scores:

```

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-184-614-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrPheLeuYsAlaArgLeuIleCySer 311

```
RESULT 74 f
US-10-184-623-348
; Sequence 348, Application US/10184623
; Publication No. US200300322129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C210
; CURRENT APPLICATION NUMBER: US/10/184,623
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-623-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-184-623-348 (1-777)

QY 1040 ATTAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystiprThrPheleuylsAlaArgLeuileCysSer 311

RESULT 75
US-10-184-635-348
; Sequence 348, Application US/10184635
; Publication No. US20030032130A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C215
; CURRENT APPLICATION NUMBER: US/10/184,635
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-635-348
```

```
Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-184-635-348 (1-777)

QY 1040 ATTAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystiprThrPheleuylsAlaArgLeuileCysSer 311

Search completed: August 3, 2003, 10:58:20
Job time : 124 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: August 3, 2003, 10:29:36 ; Search time 401 Seconds

(without alignments)
11759.598 Million cell updates/sec

Title: US-09-774-490-1
Sequence: 1 aactcttatttcacgagc.....agccttttcccaataacc 2709

Scoring table: OLIGO

Xgapop 60.0, Xgapext 60.0
Fgapop 60.0, Fgapext 60.0
Delop 6.0, Delext 7.0

Searched: 5580241 seqs, 870357830 residues

Word size: 1

Total number of hits satisfying chosen parameters: 10777494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 100 summaries

Command line parameters:

MODEL=frame+n2p.model -DEV=xlh
O=/cgn2_1/USPTO/US09774490/runat_03082003.102928.4506/app.query.fasta_1.2887
-DB=Pending_Patents_AA_Main -QMT=fastan -SUFFIX=oligo.rapm -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=BITS -START=1 -END=-1 -MATRIX=oligo
-ALIGN=human40.cdi -LIST=100 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-MAXLEN=200000000 -USER=US09774490 @cgn 1.1 628 @runat_03082003.102928.4506
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=50 -WAIT -DSPBLOCK=100
-LONGLOG -DB TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*

29: /cgn2_6/ptodata/2/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	771	86.1	771	1 PCT-US00-41943-24	Sequence 24, Appl
2	771	86.1	771	1 PCT-US02-07826-284	Sequence 284, Appl
3	771	86.1	771	8 US-08-488-935-2	Sequence 2, Appl
4	771	86.1	771	8 US-08-488-057-2	Sequence 2, Appl
5	771	86.1	771	20 US-09-694-085-24	Sequence 24, Appl
6	771	86.1	771	22 US-09-791-537-31106	Sequence 31106, A
7	771	86.1	771	26 US-10-067-632-54	Sequence 54, Appl
8	771	86.1	771	26 US-10-097-340-284	Sequence 284, Appl
9	771	86.1	771	28 US-10-247-671-164	Sequence 164, Appl
10	771	86.1	771	28 US-10-262-538-10	Sequence 10, Appl
11	771	86.1	771	31 US-06-323-784-164	Sequence 164, Appl
12	287	32.1	287	21 US-09-791-537-32993	Sequence 32993, A
13	270	30.2	272	21 US-09-724-676-66689	Sequence 66689, A
14	270	30.2	272	21 US-09-724-676A-66689	Sequence 66689, A
15	151	16.9	151	23 US-09-864-761-47112	Sequence 47112, A
16	151	16.9	151	28 US-10-203-135-32555	Sequence 32555, A
17	151	16.9	151	28 US-10-203-136-33448	Sequence 33448, A
18	92	10.3	772	22 US-09-791-537-72548	Sequence 72548, A
19	92	10.3	772	22 US-09-791-537-101922	Sequence 101922, A
20	82	9.2	666	22 US-09-791-537-62757	Sequence 62757, A
21	79	8.8	772	16 US-09-282-167-32	Sequence 32, Appl
22	79	8.8	772	22 US-09-791-537-128280	Sequence 128280, A
23	73	8.2	76	31 US-06-160-203-3266	Sequence 3266, Ap
24	73	8.2	76	31 US-06-169-840-4970	Sequence 4970, Ap
25	58	6.5	772	22 US-09-791-537-111919	Sequence 111919, A
26	57	6.4	76	31 US-06-181-996-1113	Sequence 1113, Ap
27	54	6.0	54	26 US-10-029-186-30748	Sequence 30748, A
28	53	5.9	57	1 PCT-US01-00663-33648	Sequence 33648, A
29	53	5.9	57	23 US-09-864-761-44075	Sequence 44075, A
30	53	5.9	57	27 US-10-182-993-32519	Sequence 32519, A
31	53	5.9	57	28 US-10-203-134-33356	Sequence 33356, A
32	53	5.9	57	28 US-10-203-135-32559	Sequence 32559, A
33	53	5.9	57	28 US-10-203-136-33412	Sequence 33412, A
34	53	5.9	57	28 US-10-203-137-33541	Sequence 33541, A
35	53	5.9	57	28 US-10-203-139-32331	Sequence 23331, A
36	53	5.9	60	26 US-10-029-186-29372	Sequence 29372, A
37	53	5.9	101	19 US-09-500-746-13	Sequence 13, Appl
38	48	5.4	50	31 US-09-500-746-13	Sequence 3558, Ap
39	48	5.4	50	31 US-06-160-203-3958	Sequence 4925, Ap
40	41	4.6	41	23 US-09-864-761-44552	Sequence 44552, Ap
41	41	4.6	41	27 US-10-182-993-33490	Sequence 33490, A
42	41	4.6	41	28 US-10-203-135-33728	Sequence 33728, A
43	41	4.6	41	28 US-10-203-136-34569	Sequence 34569, A
44	38	4.2	53	31 US-06-177-571-3274	Sequence 3274, Ap
45	35	3.9	507	19 US-09-500-746-9	Sequence 9, Appl
46	33	3.7	33	1 PCT-US01-00663-33648	Sequence 33648, A
47	33	3.7	33	23 US-09-864-761-39708	Sequence 39708, A
48	33	3.7	33	27 US-10-182-993-32618	Sequence 32618, A
49	33	3.7	33	27 US-10-182-995-26180	Sequence 26180, A
50	33	3.7	33	28 US-10-203-134-33472	Sequence 33472, A
51	33	3.7	33	28 US-10-203-135-32678	Sequence 32678, A
52	33	3.7	33	28 US-10-203-136-33571	Sequence 33571, A
53	33	3.7	33	28 US-10-203-137-33648	Sequence 33648, A
54	33	3.7	33	28 US-10-203-137-32427	Sequence 32427, A
55	33	3.7	50	31 US-06-177-571-3503	Sequence 3503, Ap
56	33	3.7	60	31 US-06-182-467-1352	Sequence 1352, Ap
57	33	3.7	60	31 US-06-182-467-1357	Sequence 1357, Ap
58	31	3.5	35	31 US-06-183-123-1597	Sequence 1597, Ap
59	24	2.7	778	22 US-09-791-537-105918	Sequence 105918, A
60	22	2.5	67	19 US-09-513-999C-6632	Sequence 6632, Ap

61 17 1.9 17 1 PCT-US00-41943-16 Sequence 16, Appl
62 17 1.9 17 1 PCT-US00-41943-43 Sequence 43, Appl
63 17 1.9 17 20 US-09-694-085-16 Sequence 16, Appl
64 17 1.9 17 20 US-09-694-085-43 Sequence 43, Appl
65 17 1.9 73 31 US-60-177-571-3765 Sequence 3765, Ap
66 17 1.9 73 31 US-60-177-646-3105 Sequence 3105, Ap
67 17 1.9 79 31 US-60-177-571-3766 Sequence 3766, Ap
68 17 1.9 779 18 PCT-US00-31116-2 Sequence 2, Appl
69 17 1.9 779 18 US-09-455-560-2 Sequence 2, Appl
70 17 1.9 782 21 PCT-US00-04340-18 Sequence 18, Appl
71 17 1.9 782 19 US-09-507-209-18 Sequence 18, Appl
72 17 1.9 782 23 US-09-813-290-4 Sequence 4, Appl
73 17 1.9 785 1 PCT-US00-04340-92 Sequence 92, Appl
74 17 1.9 785 19 US-09-507-209-92 Sequence 92, Appl
75 17 1.9 875 23 US-09-813-290-2 Sequence 2, Appl
76 17 1.8 17 1 PCT-US00-41943-2 Sequence 2, Appl
77 16 1.8 17 20 US-09-694-085-2 Sequence 2, Appl
78 16 1.8 34 31 US-60-177-571-3504 Sequence 3504, Ap
79 16 1.8 764 22 US-09-791-537-85681 Sequence 85681, A
80 16 1.7 39 31 US-60-170-373-3437 Sequence 3437, Ap
81 15 1.7 89 31 US-60-196-718-4235 Sequence 4235, Ap
82 15 1.7 110 31 US-60-177-571-4486 Sequence 4486, Ap
83 15 1.7 478 1 PCT-US02-40059-15 Sequence 15, Appl
84 15 1.7 539 22 US-09-760-479-723 Sequence 723, App
85 15 1.7 539 28 US-10-206-008-723 Sequence 723, App
86 15 1.7 539 28 US-10-211-364-1199 Sequence 1199, Ap
87 15 1.7 540 22 US-09-760-479-483 Sequence 483, App
88 15 1.7 540 28 US-10-206-008-483 Sequence 483, App
89 15 1.7 636 31 US-60-389-987-2003 Sequence 2003, Ap
90 15 1.7 636 31 US-60-412-418-2003 Sequence 2003, Ap
91 15 1.7 658 31 US-60-242-679-840 Sequence 840, App
92 15 1.7 691 16 US-09-262-167-33 Sequence 33, Appl
93 15 1.7 701 16 US-09-262-167-29 Sequence 29, Appl
94 15 1.7 749 1 PCT-US00-41943-25 Sequence 25, Appl
95 15 1.7 749 20 US-09-694-085-25 Sequence 25, Appl
96 15 1.7 749 28 US-10-262-538-12 Sequence 12, Appl
97 15 1.7 749 28 US-10-285-351B-1 Sequence 1, Appl
98 15 1.7 749 31 US-60-335-783-1 Sequence 4, Appl
99 15 1.7 751 1 PCT-US00-10046-4 Sequence 4, Appl
100 15 1.7 751 1 PCT-US00-10046-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
PCT-US00-41943-24
Sequence 24, Application PC/TUS0041943
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
FILE REFERENCE: 00786-441M01
CURRENT APPLICATION NUMBER: PCT/US00/41943
CURRENT FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: US 09/694, 085
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/164, 056
PRIOR FILING DATE: 1999-11-08
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 771
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-41943-24

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-24 (1-771)
QY 200 ATGGGCTGGTTAACTGATGATTTCTGCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
1 MetGlyTyrLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrIleArgAla 20
QY 260 AACTGATCAGAAATGGGAGAGAACAAATGTCGCCAGGCTGAAATTATCTTACAAAGAAATGTTG 319
21 AenTYrGlnAsnGlyValAsnAsnValProAlaGlyLeuValSerTyrIleGlyMetLeu 40
QY 320 GAATCAACAATGTATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
41 GluSerAsnAsnValIleThrPheAsnGlyLeuIleAsnSerSerTyrIleThrPhe 60
QY 380 CTTTGGATGAGGAGGAGGAGTGGCTGTATGTTGGAGCAAGAGACATATTTTCATTC 439
61 LeuLeuAspGlyGlnArgSerArgLeuValGlyAlaValAspHisIlePheSerPhe 80
QY 440 GACCTGTTAATATATCAAGATTTTCAAAAGATTGGTGGCCAGTATCTTACACAGAGA 499
81 AspLeuValAsnIleValAspPheGlnIleValIleProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGGCTGGAAAAGACATCTGGAAGAATGTCTTAATTTCATCAAGTA 559
101 AspGlyCysGlyTrpAlaGlyValAspIleLeuValGlyCysAlaAsnPheIleVal 120
QY 560 CTTAAGGATTAATACAGACTCTGTAAGCTGTGGAACGGGGGCTTTTCATCAATT 619
121 LeuValAlaTyrAsnGlnThrIleLeuValAlaCysGlyThrGlyValAspHisProIle 140
QY 620 TGCACCTCATGTAATTTGACATCATCTGAGAGCAATATTTTAACTGGAGAACTCA 679
141 CysThrTyrIleGlnIleGlyHisIleHisProGluAspAsnIlePheValGlyLeuSer 160
QY 680 CATTTTGAAGACGGCGGTGGAGAGATCCATATGACCTTAAGCTGTACAGATCCCTT 739
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrIleSerLeu 180
QY 740 TTAATAGATGAGATTTATCTGTGAGCTGAGCTGATTTTAAAGGCGGAGACTTCT 799
181 LeuIleAspGlyGlnLeuTyrSerIleThrAlaAlaAspPheMetGlyValArgAspPheAla 200
QY 800 ATCTTCGAACCTCTGGGACACCAACCAATTCAGAGACAGACATGATTCAGGTGG 859
201 IlePheArgThrLeuGlyHisIleHisProIleArgThrGlnGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTATATGATGTCACCTCATCTCAAGAGTGAATCTTGAAGAT 919
221 LeuAsnAspProLysPheIleSerIleHisIleLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTCTCCGTGGAATGCAATGAGATGGAACACTCTGGAAGACT 979
241 AspLysValIleTyrPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyValAla 260
QY 980 ACTCAGCTTAAGTATGATGATGATGCAATGCAAGATGCACTTGGAGGGCAGACAGATCTGGTG 1039
261 ThrHisAlaArgIleGlyGlnIleValValAsnAsnAspPheGlyGlnHisArgSerIleVal 280
QY 1040 AATAAATGACAACTTCTCAAGCTGTGATTTCTCAAGTCCAGTCCAAATGTC 1099
281 AsnLysTyrPheThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACATCTTTTGAAGACTGAGAGATGATATTCCTTAAGAACTTAAAGATCTTAA 1159
301 IleAspThrHisPheAspGlyLeuGlnAspValPheLeuMetAsnPheValAspProLys 320
QY 1160 AATCAGTTGATATGAGTGTATTACGATTCAGTTCAGTAACATTTTCAAGAGATGACCCGTG 1219
321 AsnProValValIleTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATATGATGATGATGATGAGAAAGGATGTTCTTGGTCCATATGATCCCAAGGAT 1279
341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCACTATCAATGAGTGGCTTATCAAGAAAGATCCCTATCCAGGCGAGAACT 1339
 DB 361 GTProbandTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAACATTGGTGGTTTGAAGTCTCAAGAAAGACCTTCCTGATGATGTTATA 1399
 DB 381 CysProSerTyrThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTGGCAAGATCATCAGCATGTATCCAGTGTCTTCCATGAAGAACATGGCCCA 1459
 DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnMetArgPro 420
 QY 1460 ATAGTATCAAAACCGATGTAATTAATTATCAATTTTACAAATTTGCTAGAACCGATGAT 1519
 DB 421 IleValIleIleTyrThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAAGATGACAGATGATGATGTTATGTTATCGAAACGATGTTGGACGTTCTTAA 1579
 DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGAGACTGTGATGTTATGAAGAGATGCTGTGGAAGAAAG 1639
 DB 461 ValValSerIleProLysGlnThrTyrTyrAspLeuGluValLeuLeuGluLysMet 480
 QY 1640 ACAGTTTTCGGGAAACCGACTGCTATTTTCAAGCAATGAGACTTCCACTAAGCAGCAACA 1699
 DB 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGATTCAACGGCTGGGGTGGCCAGCTCCCTTACACCGATGATATTATAC 1759
 DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGGAAAGCGTGTGCTGAGTGGTCCCTCGCCGAGAACCTTACTGTGCTGGAGTGTCT 1819
 DB 521 GlyLysAlaCysAlaGlyCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
 QY 1820 GCATGTTTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATTAAGAAT 1879
 DB 541 AlaCysSerArgTyrPheProThrAlaLysArgGlnThrArgGlnAspIleArgAsn 560
 QY 1880 GGAAGACCACTGACTGACTGTTTCAAGCTTACACCATGATTAATCACCATGGCCAGCCCT 1939
 DB 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisIleGlyHisSerPro 580
 QY 1940 GAAAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATTTTGGAAATGCGAGTCCGAG 1999
 DB 581 GlnGluArgGlnIleIleTyrGlyValGluAsnSerSerThrPheLeuGlnCysSerProLys 600
 QY 2000 TGGCAGAGAGCGCTGCTATTTGGCAATCCAGAGGCGAAATGAAGAGCGAAAGAGAG 2059
 DB 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAlaGlnGlnGlnGlnGlnGln 620
 QY 2060 ATCAGAGTGAATCATCATCATCATCAGACAGATCAAGGCTTCTGTAAGTATGATCAAA 2119
 DB 621 IleArgValAspAspHisIleIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAAGATTGCGCAATTAACCTTGGCCATGCGGTGGAAATGGGTTTCAATAACTCTT 2179
 DB 641 GlnLysAspSerCysLysAsnTyrLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGGTAAACCTGGAAGTCAATTGACAGAGCAATTTGGAAGAACTTTTCAATAAGAT 2239
 DB 661 LeuLysValThrLeuGlnValIleIleAspThrGlnHisSleuGlnGlnLeuLeuHisLysAsp 680
 QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCCATATGATGATGACCTTACCGCAGAG 2299
 DB 681 AspAspGlyAspGlySerTyrThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGTATAGAGATTTATGAGCTGATCAACACCCCAATCTCAACGATGATGATGAG 2359
 DB 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720

QY 2360 TTTGTGAAACAATTGGAAAAAGGACCGAAACACAGTGGCAAGGCGACAGATACC 2419
 DB 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
 QY 2420 CCAGGGAACAGTAAACAATTGAAAGCATTTACAGAAATTAAGAAAGTGAACGAGAG 2479
 DB 741 ProGlyAsnSerAsnLysIleTyrHisSleuGlnGlnLeuLysValGlyArgAsnArgArg 760
 QY 2480 ACCCAGAATTGAGAGGCGACCGAGAGTGTG 2512
 DB 761 ThrHisGluPheGluArgAlaProArgSerVal 771
 RESULT 2
 PCT-US02-07826-284
 ; Sequence 284, Application: PC/TUS0207826
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
 ; FILE REFERENCE: MRI-030PC
 ; CURRENT APPLICATION NUMBER: PCT/US02/07826
 ; PRIOR FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 60/276,025
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/325,149
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: 60/276,026
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/324,967
 ; PRIOR FILING DATE: 2001/09/26
 ; PRIOR APPLICATION NUMBER: 60/311,732
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/325,102
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/323,580
 ; PRIOR FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 284
 ; LENGTH: 771
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-07826-284
 Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x PCT-US02-07826-284 (1-771)
 QY 200 ATGGCTGCTTAACTAGAGATGCTGCTTTCTGGGAGTATTAATTACTTAACAGAGAGA 259
 DB 1 MetGlyTyrPheThrThrAlaIleValCysLeuPheThrGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGATGGAGAGAACAAATGTGCCAAGCTGAATAATTACTTACAAAGAAATGTTG 319
 DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGlnMetLeu 40
 QY 320 GAATTCACATGTGATCATCTTCAATGGCTTGGCCACAGTCCAGTATTAATTAATCTTC 379
 DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGAGAGAGAGTAGGCTGTATGTTGAGAGCAAGATCATATTTTCATTC 439
 DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGTTAATATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTACACGAGAGA 499
 DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArgArg 100

QY	500	GATGTGAATGCAACTGGGCGTGGAAAAGACATCCCTGAAAAGATATGGCTAAATTTCATTCAGAGTA	555
DB	101	AspGluCysLysTrpIleGlyLysAspIleLeuLysGluCysAlaAsnPhelIeLysVal	120
QY	560	CTTAAGCATATATATCAGACTCATCTGTATACGGCTGTGGAAACGGGGCTTTTATCCCAATT	619
DB	121	LeuLysAlaTyrAsnGlnThrIshIeuleuTyrAlaCysGlyThrGlyAlaPheHisProIle	140
QY	620	TGCACCTACATTTGAAATTGGACATCATCTGTAGGACAAATTTTAACTGGAGAACTCA	679
DB	141	CysTrpTyrIleGlnIleGlyIshIshsProGluAspAsnIlePheLysLeuGluAsnSer	160
QY	680	CATTTTGGAAAACGGCCGTGGAAAGATCCATATGACCCCTAAGCTGTGACAGACTCCCTT	739
DB	161	HisPheGluAsnGlyAlaArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerIeu	180
QY	740	TTAATATAGATGAGAAATTATACTCTGGAACTGCAAGCTGATTTTATGGGGCGAGACTTTGCT	799
DB	181	LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyAlaArgPheAla	200
QY	800	ATCTTCCGAACCTTGGGACCAACCAACCAATCAGACAGACACATGATCCAGGTGG	859
DB	201	IlePheAspThrLeuGlyIshIshsProIleAspThrGlnGlnHisAspSerArgTrp	220
QY	860	CTCAATGATCCAAAGCTTCAATTAGTCCCACTCATCTCAGAGAGTGCACATCTTCAAAAT	919
DB	221	LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp	240
QY	920	GACAAAGATTAATCTTTTCTTCCGTGAAAATGCATATGATGGAGAACCTCTGGAAAGCT	979
DB	241	AspLysValTyrPhePhePheAspGluLysAlaIleAspGlyGlnHisSerGlyLysAla	260
QY	980	ACTCAGCTAGAAATAGTGCAGATATGACAAATGCACTTTGGAGGGCAGACAGAGTGTGTG	1039
DB	261	ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyIshIshArgSerIeuVal	280
QY	1040	AATAAATGGACAAACATTCTCCAAAGCTGTCTGATTTGCTCAGTGCACAGTCCAAATGAC	1099
DB	281	AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly	300
QY	1100	ATTGACACTCATTTTGGATGAACTGGAGAGTGAATTCCTAATGAACTTTAAAGATCCTTAA	1159
DB	301	IleAspThrHisAspHisAspIleuGlnAspValPheLeuMetAsnPheLysAspProLys	320
QY	1160	AATCCAGTGTATATGAGGTGTTTACGACTTCAGATAATTTTCAAGGATTCAGCCGTG	1219
DB	321	AsnProValIleTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal	340
QY	1220	TGTATGTATAGATAGATGATGTGGAAGGGTGTCTTGGTTCATATGGCCACAGGGAT	1279
DB	341	CysMetTyrSerMetSerAspValAlaArgValPheLeuGlyProTyrAlaHisArgAsp	360
QY	1280	GGACCACTATCATGGGTGCCTTATCAAGAAAGATCCCTATCCACGGCCAGAGAACT	1339
DB	361	GlyProAsnTyrGlnTrpValProTyrGlnGlnArgValProTyrProAspProGlyThr	380
QY	1340	TGTCCAGCAAAACATTTGGTGGTTTGACTCTACAAAGACCTTCTGATGATGTATA	1399
DB	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
QY	1400	ACCTTGGAAAGATCCACAGCCATGATCAATCCAGGTTCCTATGAAACATGAGCCCA	1459
DB	401	ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetCAsnAsnArgPro	420
QY	1460	ATAGTATCAAAAACGATGTAAATTTATCATTTTACAAAATTTGTCTGAGACCGAGTGAT	1519
DB	421	IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp	440
QY	1520	GCAGAAAGATGACAGTATGATGTATATGTTTATCGGAACAGATGTGGACCGTCTTATA	1579
DB	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460

QY	1580	TTAGTTTCATTTCCAAAGGACACTGGTAATATTAGAAAGGTTTCGTGGAAAGATG	1639
Db	461	ValValSerLLeprPolysgluThrTrpZraspLeuGluValLeuLeuGluGluMet	480
QY	1640	ACAGTTTTCCGGAAACCGACTGTCTATTTCAGCAATGAGAGCTTTCACATAAGCAACA	1699
Db	481	ThrValPheArgGluProThrAlaIleSerIleMetGluLeuSerThrLysGlnGln	500
QY	1700	CTATTATATTGGTTCAAGGGCTGGGGGTTGCCAGCTCCCTTTTACACCGGTGATATTTC	1759
Db	501	LeuTyxIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyx	520
QY	1760	GGGAAAGCGTGTGCGAGTGTGGCTGGCCCGCAACCCCTTACCTGTGCTGGGAAGTCT	1819
Db	521	GlyLysAlaCysValaGluCysCysLeuAlaArgAspProTyxCysAlaIrrAspGlySer	540
QY	1820	GCATGTTCTCGCTATTTTCCCACTGGCAAGAAGACGCAAGAAGCAAGATATTAAAGAT	1879
Db	541	AlaCysSerAsnArgTyxPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn	560
QY	1880	GGAGACCCACTGACTACTGTTCAGACTTACACCATGATATATCCATGGCCACAGCCTT	1939
Db	561	GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro	580
QY	1940	GAAGAAGAAATCATCTATGTGTGTAAGAAATATATATGACACTTTTGGATATGACATGCCAAG	1999
Db	581	GluGluArgIleIleTyxGlyValaGlnAsnSerSerThrPheLeuGluCysSerProLys	600
QY	2000	TCGACAGAGCGCTGTCTATTGGCAATTCACAGCGCGAAATGAAAGACCAAAAGAGAG	2059
Db	601	SerIleThrArgAlaLeuValTyxTrpGlnPheGlnArgArgAsnGluValaGlyGluGln	620
QY	2060	ATCAGAGTGATGATCATATCATATGACAGACAGATCAAGGCTTCTGTACGTAGTCTACAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln	640
QY	2120	CAGAAAGATTCAGGCAATTCCTCTGCGCAATGCGGTGGAACAATGGGCTTCATACAACTCT	2179
Db	641	GlnLysAspSerGlyAsnTyxLeuCysHisAlaValaGlnHisGlyPheIleGlnThrLeu	660
QY	2180	CTTAAGGTAAACCCCTGGAAGTCATTGACACAGAGACTTTGGAAGAACTTCTTCATTAAGAT	2239
Db	661	LeuLysValaThrLeuGluValaIleAspThrGlnHisLeuGluGluGluLeuLeuHisLysAsp	680
QY	2240	GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATAGACTGACACCTTACGACAG	2299
Db	681	AspAspArgLysAspArgLysSerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys	700
QY	2300	GTCGTGTATCAGAGACTTCATGACAGCTTATCAACACCCCAATCTCAACAAGATGATGAG	2359
Db	701	ValIrrTyxIrrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
QY	2360	TTCTGTGTAAACAGTTTGGAAGAGGACCGAAAAACAAGTCGGGCAAGAGCGACAGACATACC	2419
Db	721	PheCysGluGlnValaIrrLysArgAspArgLysGlnArgGlnArgGlnArgProGlnHisThr	740
QY	2420	CCAGGGAACAGTAAACAATGGAAGCACTTCAAGAAATATAGAAAGGTAGAAAACAGAGG	2479
Db	741	ProGluAsnSerAsnLysIrrTyxHisLeuGlnGlnLysAsnLysLysGluArgAsnArgArg	760
QY	2480	ACCAACGAATTTGAGAGGGACCCACGAGAGTGC	2512
Db	761	ThrIleGluPheGluArgAlaProLysSerVal	771

RESULT 3
US-08-488-935-2
Sequence 2 Application US/08488935-2
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy

```

TITLE OF INVENTION: Semaphorins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESS: RICHARD ARON OSMAN, Ph.D.
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,935
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/489,057
FILING DATE: 09-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-2/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-935-2

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-08-488-935-2 (1-771)
QY 200 ATGGCTGTTAACTAGAGTTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
DB 1 MetGlyTTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTTCGAAATGGGAAGAAACAATGTCGCAAGCTGAATTTATCCACAAAGAAATGTTG 319
DB 21 AenTYrgInAsnGlyAsaAsnValProArgLeuValLeuSerTyrIysGluMetLeu 40
QY 320 GAATCAACAATGATGATCACTTCAATGAGCTTGCCACAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGAGTGGCTGTATGTTGAGCAAGAGATCATATTTTCATTTC 439
DB 61 LeuLeuAspGluGluLysSerArgLeuTyrValGlyAlaAspHisIlePheSerPhe 80
QY 440 GACCTGTTAAATATCAAGATTTTCAAAAGATTTGTGCGCATATCTTACACCGAAGA 499
DB 81 AspLeuValAsnIleLysAspPheGlnTyrIleValTrrProValSerTyrThrArgArg 100
QY 500 GATGATGCAAGTGGCTGGAAGAAAGACATCCGAAAGATGCTATTTTCAATCAAGATA 559
DB 101 AspGluCysIleTrrPheIleGlyAspIleLeuLysGluCysAlaAsnHeIleLysVal 120
QY 560 CTTAAAGCATATATATCAAGCTGTTGACGCTGTGGAACGGGGCTTTTCATCCAAAT 619

```

```

DB 121 LeuValAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTCATTTGAATTTGACATCATCTCGAGCAATATTTTATAGCTGAGAACCTCA 679
DB 141 CysThrTyrIleGlnIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTCGAAAACGGCGTGGGAAGATCATATGACCTTGAAGCTGTCAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrIleSerLeu 180
QY 740 TTAATAGATGAGAAATTAATCTGTGAACCTGCACTGATTTTATGGCGGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTCTGGGACACCAACCAATCAGAGACAGACATGATTCACAGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisSerArgTrr 220
QY 860 CTCATATGATCCAAAGTTATAGTCCACCTCATCTCAGAGAGTACATCTGAAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTCCGTGAAGAAATGCAATGATGAGAACTCTGGAAGCT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATGTCGATGATGCAAGATGATGCTTGGAGGGCAGCAAGTCTGGT 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACAAACATTTCTCAAGCTGCTGATTTGCTCAGTCCAGGTCCAATGGC 1099
DB 281 AenLysTrrPheThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTGATGAACTGCAAGATGATTTCTTAATGACATTTTAAATCCTTAA 1159
DB 301 IleAspThrHisAspAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGATATGAGAGTGTTCAGCTTCCATTAACATTTTGAAGGAACACCGG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTTATGATGATGATGATGAGAGGGGTGCTTGTGCTATGATGCCACAGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGAGGTGCTTATCAAGGAAGATCCCTATCAACGGCCAGAACT 1339
DB 361 GlyProAsnTyrGlnTrrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCACGCAAAACATTTGGTGGTGTTCATGCTTAAAGAGACCTTCGAGAGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGAAGAAGTCATCCAGCCATGATACATCCAGTGTTCATGATGAACATGCCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGATGTAATTAATTAATTAACAAATTTGTAACCGAGGTGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp 440
QY 1520 CGAGAAGATGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTTCTTAA 1579
DB 441 AlaGluAspGlyGlnTrrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTTCCAAAGAGACTTGATGATTTTGAAGAGGTTCTGCTGGAAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTGGGAAACGACTGCTTATTTCAAGATGAGGCTTTTCACTAAGCAGCAAA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

```

```

QY 1700 CTATATATGTTCAACGGCTGGGCTTGCCAGCTCCCTTAAACCGGCTGATATATAC 1759
    |||||
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGGATGTCTT 1819
    |||||
Db 521 GlyValAlaCysAlaGlnCysCysLeuAlaArgAspProTyrCysAlaIleTyrAspGlySer 540
QY 1820 GCATGTTCTGGCTATTTTCCCATCTGCAAGAAGACGCAAGAGAGACAGATATTAAGAAAT 1879
    |||||
Db 541 AlaCysSerArgTyrPheProThrAlaValArgArgThrArgGlnArgIleArgAsn 560
QY 1880 GGAGACCACTGACTCACTGTTCAGACTTACACCATGATATATACCATGCGCACAGCCCT 1939
    |||||
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTTATGGGTGTAAGAGATATGACATATTTTGAATCCAGTCCGAGAG 1999
    |||||
Db 581 GluGlnAlaGlyIleIleTyrGlyValGluAsnSerSerThrPheLeuGlnCysSerProLys 600
QY 2000 TGCCAGAGAGCGCTGTCTTGTGGCAATTCAGAGCGGAATGAAAGCGAAAGAGAG 2059
    |||||
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgGlnArgGlnGlnGlnGlnGln 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTGTCTAGCTAGTCTTAA 2119
    |||||
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnIleuLeuLeuAspSerLeuGln 640
QY 2120 CAGAGAGATTCAGAGCAATTAACCTCTGCGCATGCGGTGAGACATGAGTTCATCAAACTTT 2179
    |||||
Db 641 GlnLysAspSerGlyAsnTyrLeuGlyHisAlaValGlnIleGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTTGAAGTCATGACAGACATTTGGAAGAACTTCTCTTAAAGAT 2239
    |||||
Db 661 LeuLysValThrLeuGlnValIleAspThrGlnHisLeuGlnIleuLeuHisLysAsp 680
QY 2240 GATGATGAGATGAGCTCTAAGACCAAGAAATGTCATAGTACACCTTACCCAGAGAG 2299
    |||||
Db 681 AspAspGlyAspIleTyrSerThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGATCAGAGATTCATGACAGCTTCATCAACCCCAATTCATCAACAGATGATGAG 2359
    |||||
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAAGTTTGAAAGAGGACCAAGAAACAAGTCGCGAAAGCGCAGACATACC 2419
    |||||
Db 721 PheCysGlnGlnValIleTyrLysArgAspArgLysGlnArgGlnArgProGlyHisIleThr 740
QY 2420 CCAGGAGACATTAACAATGSAAGCACTTACAAGAAATTAAGAAAGTAGAAACAGAGAG 2479
    |||||
Db 741 ProGlyAsnSerAsnLysIleTyrHisIleuGlnGlnLysAsnLysLysIleArgAsnAlaGly 760
QY 2480 ACCCAGCAATTTGAGAGGCGACCCAGAGAGTTC 2512
    |||||
Db 761 ThrHisGlnPheGlnArgAlaProArgSerVal 771

```

```

RESULT 4
US-08-489-057-2
; Sequence 2, Application US/08489057
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA

```

```

; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/489,057
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/BAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-489-057-2
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
Gaps: 0
DB:
US-09-774-490-1 (1-2709) x US-08-489-057-2 (1-771)
QY 200 ATGGCTGGTTTAATCAGATTTGCTGTCTTTTCTGGGAGTATTAATTAACAGCAAGAGA 259
    |||||
Db 1 MetGlyTrpLeuThrThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGAAGACATATGTCACAGCTGGAATTAATCTTACAAAGAAATGTTG 319
    |||||
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrIleGlnMetLeu 40
QY 320 GAATCAACAATGATCATCTTCAATGCTTGCGCAACAGCTCCAGTATCATACCTTC 379
    |||||
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGAGACGAGTAGAGCTGTATGTTGGAGCAAGATCACATATTTTCATTC 439
    |||||
Db 61 LeuLeuAspGlnGlnLysSerThrArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTATATATCAAGGATTTTCAAAAAGATTTGTGGCCAGTATCTTACACAGAAAG 499
    |||||
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValIleProValSerTyrThrArgArg 100
QY 500 GATGAATGCAATGGGCTGGAAGAAACAATCCCTGGAAGAAGATGTCTAATTTCAAGGTA 559
    |||||
Db 101 AspGlnCysLysTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGCATATTAATCAAGACTCACTTACGCGCTGTGAAACGGGGCTTTTCATCCAAAT 619
    |||||
Db 121 LeuLysAlaIleTyrAsnGlnThrHisIleuLysTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTTACATTAATGAATGACATCATCTGAGAGCAATATTTTAACCTGGAGAACTCA 679
    |||||
Db 141 CysThrTyrIleGlnIleGlyHisHisProGlnAspAsnIlePheLysLeuGlnLysSer 160
QY 680 CATTTTGAAGAGCGCGGTGGAGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTT 739
    |||||
Db 161 HisPheGlnAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180

```

```

QY 740 TTAATGATGAGAAATTATCTGGAAGCTGAGCTGATTTATGAGGCGAGACTTGGT 799
Db 181 LeuileaspGlyGluLeuLeuSerGlyThrAlaAlaAspPheMetCylArgAspPheAla 200
QY 800 ATCTCCGAACTCTTGGGACCAACCAATCAGAGACAGAGATGATATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyAlaShiShiAspProIleArgThrGluGlnIleAspSerArgTyr 220
QY 860 CTCGAATGATCCAAAGTTCATTAATGAGCCCACTCATCTCAGAGATGAGACATCTGAAGAT 919
Db 221 LeuAlaAspProLeuPheIleSerAlaShiLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTCTTCCGTGAAAATGCAATAGATGAGAACACTCTGAAAAGCT 979
Db 241 AspIleValTyrPhePhePheArgGluAlaIleAspGlyGluIleSerGlyAla 260
QY 980 ACTCAGCTGAGAAATGCTGATGATGCAAGATGATGATGAGGCGACAGAAAGTCTGGTG 1039
Db 261 ThrAlaAlaArgGlyGlyGlyIleCysIleAspAspPheGlyGlyIleArgSerLeuVal 280
QY 1040 AATAATGAGACAAACATCTCTCAAGCTGCTGATGATGCTGATGCTGAGCTGCAATGAGC 1099
Db 281 AsnIleTyrPheThrPheLeuIleAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATGCACTCATTTTATGATGAGTGAAGTGAAGTATTCCTATGATGATGATGATGATGATGAT 1159
Db 301 IleAspThrIleAspAspGlyLeuGluIleAspValPheLeuMetAsnPheIleAspProIys 320
QY 1160 AATCCAGTTGATGATGAGTGTTCAGCTTCCAGTACATTTTCAAGGAGTCAAGCTG 1219
Db 321 AsnProValIleValTyrGlyValPheThrThrSerSerAsnIlePheIleGlySerAlaVal 340
QY 1220 TGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
Db 341 CysMetTyrSerMetSerAspValAlaArgValAlaPheLeuGlyProIleAlaIleArgAsp 360
QY 1280 GAGACCAACTATCAATGAGTGTGCTTATCAAGAGAGAGTCCCTATCCAGCGCAGAACT 1339
Db 361 GlyProAsnTyrGlnTyrValProIleArgIleValAlaValProIleProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGTGTTTGAAGTCAACAAAGAGCTTCTGATGATGATGAT 1399
Db 381 CysProSerLeuTyrPheGlyGlyIlePheAspSerThrIleAspLeuProAspValIle 400
QY 1400 ACCTTTGAAGAGATCATCCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
Db 401 ThrPheAlaArgSerHisProAlaMetCylArgAspProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACCGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db 421 IleValIleIleValThrAspValAlaAsnTyrGlnPheThrGlnIleValAlaAspArgValAsp 440
QY 1520 GCAGAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuIleVal 460
QY 1580 GTAGTTGAATTCCTCAAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1639
Db 461 ValIleValSerIleLeuIleValThrIleValThrIleValThrIleValThrIleValThr 480
QY 1640 ACAGTTTTTCCGGAACCGATGCTGATTTTCAAGATGAGCTTCCAGTATGAGAGAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrIleGlnIleGln 500
QY 1700 CTATATATTTGTTCAACGGCTGGGGTGGCCAGCTCCCTTACACCGGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyAlaAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAAGGCTGTGAGTGTGCTGCGCGCGAGAGCCCTTATGCTGATGATGATGATGATGATGAT 1819
Db 521 GlyIleValAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIleTyrAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1879

```

```

Db 541 AlaCysSerArgTyrPheProThrAlaIleValArgAlaGlyThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGAGCCCACTGATCTCATCTGTTCAAGCTTACACCAATGATTAACACCAAGCCCAAGCCCT 1939
Db 561 GlyAspProLeuThrIleCysSerAspLeuHisShiAspAsnHisShiGlyHisSerPro 580
QY 1940 GAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1999
Db 581 GluGluAlaGlyIleIleTyrGlyAlaGluAsnSerSerThrPheLeuGluCysSerProIys 600
QY 2000 TCGCAGAGAGCGCTGCTGATTTGCAATTCAGAGCCCAATGAAAGAGAGAGAGAGAGAGAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAlaGlnGluGluArgIleGlyGluGln 620
QY 2060 ATCAGAGTGAATGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuAspSerLeuGln 640
QY 2120 CAGAGAGATTCAGCAATTAATCTGCGCATGCGGTGGAACATGGGTTCATCAAACTCTT 2179
Db 641 GlnIleAspSerGlyAsnTyrLeuCysShiAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCCCTGAAAGTCAATGACACAGAGATTTGAGAGAACTTCTTCAATAAGAT 2239
Db 661 LeuIleValThrLeuGluValIleAspThrGlnHisIleGluGluGluLeuHisIleAsp 680
QY 2240 GATGATGAGATGCTCTTCAAGACCAAGAAATGTCCATAGATGACACCTTACCCAGAG 2299
Db 681 AspAspGlyAspGlySerIleTyrThrIleGlnLeuSerAsnSerMetThrProSerGlnIys 700
QY 2300 GTCTGTACAGAGATTCATGATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAAATTGAAAAAGGACCGCAAAACACACTGCGCAAAAGCCAGACATACC 2419
Db 721 PheCysGluGlnValTyrPheArgAspArgIleGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAAGGAACTGATCAAAATGAGACATTAACAAATAAGAAAGTGAAGAGAGAGAGAGAGAG 2479
Db 741 ProGlyAsnSerAsnIleTyrPheIleGlnGlnIleGlnIleGlnIleGlnIleGlnIleGln 760
QY 2480 ACCCAAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 5
US-09-694-085-24
; Sequence 24, Application US/09694085
; GENERAL INFORMATION:
; APPLICANT: Behar, Oded
; APPLICANT: WOOLF, Clifford J.
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441001
; CURRENT APPLICATION NUMBER: US/09/694,085
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; PRIOR FILING DATE: 1999-11-08
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-694-085-24

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0

```

DB: 20 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-694-085-24 (1-771)
QY 200 ATGGGCTGGTTAACTAGATGTTCTGCTTTCTGGGGAGTATTACTTACAGCAAGCA 259
DB 1 MeGGlyTrpLeuThrTrpLeuValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTACAGAAATGGAGAAACAATGTCGCAAGGCTGAAATTAATCTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGlnMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTCAATGCGTTGGCCAAACGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrIleThrPhe 60
QY 380 CTTTGGATGAGAAAGGAGTAGGCTGTATGTTGGAGCAAAAGATCATATTTTCATTTC 439
DB 61 LeuLeuAspGlnGlnLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAAATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGCAAG 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValIleProValSerTyrThrAlaArg 100
QY 500 GATGAATGCAAGTGGGCTGGAAAAGACATCTCGAAAAGATGTCATAATTCATCAAGTA 559
DB 101 AspGlnCysLeuTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAAGACTCACTGTACGCTGTGGAAAGGGGGCTTTTCATCCAAAT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyTyrGlnAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAATTTGACATCATCCAGAGCAAAATTTTAACTGGAGCAAGTCA 679
DB 141 CysThrTyrIleGlnIleGlyHisIleProGlnAspAsnIlePheLysLeuGlnAsnSer 160
QY 680 CATTTTGAACCGCGCTGGGAAGAGTCATATGACCTTAAGCTGTGCACAGCATCCCT 739
DB 161 HisPheGlnAsnGlyLysGlyLysSerProTyrAspProLysLeuThrAlaSerLeu 180
QY 740 TTAATGATGAGAAATTAATTAATCTGGAATGCGAGCTGATTTTAATGGGGAGCACTTTC 799
DB 181 LeuIleAspGlyGlnLeuTyrSerGlyTyrThrAlaAspPheMetGlyLysAspPheAla 200
QY 800 ATCTTCGAACCTCTGGGCGACCAACCAATCAGAGCAGAGCATGTTCCAGATGG 859
DB 201 IlePheArgThrLeuGlnHisIleHisProIleArgThrGlnGlnHisAspSerArgTyr 220
QY 860 CTCATGATCCAAAGTTCAATTAAGTCCACCTCATCTCAGAGAGTACAATCCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisIleuIleSerGluSerAsnProGlnAsp 240
QY 920 GACAAAGTATACCTTTTCCGTGAAAATGCAATGATGAGAGAACACTCTGAAAAGCT 979
DB 241 AspLysValTyrPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTGACGCTAGATAGTATGAGATATGCAAGATGACTTTGGAGGGGACAGAGTCTGGTG 1039
DB 261 ThrHisAlaArgIleGlnIleGlnIleCysLysAsnAspPheGlyGlnHisArgSerLeuVal 280
QY 1040 AATAAATGACAACATCTCTCAAGCTCGTGTGATTTGCTCAGTGCAGAGTCCAAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACATCTATTTGATGAATGCGAGATGTAATTTCTAATGAACCTTAAAGATCTTAA 1159
DB 301 IleAspThrHisAspArgIleLeuGlnAspValPheLeuMetCAsnPheLysAspProLys 320
QY 1160 AATCAGTTGATATGAGAGTGTTTACGATCCAGTAACATTTTCAAGGGATCAGCCGG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGATGTGAGAAAGGTGTTCTTGATCATATGCCACAGGAT 1279

DB 341 CysMetLysSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCATCATGGTGGCTTATGCAAGAAAGATCCCTATCCAGGCGCAGAACT 1339
DB 361 GlyProAsnTyrGlnThrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGACTCTACAAAGACCTTCTGATGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGAAGAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetCysAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGCATCAAAACGAGTGTAAATTAATCAATTAACAACAAATGTGTAGACCGATGGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGATGATGTATGTTATCGGAACAGATTTGGACCGTCTTAA 1579
DB 441 AlaGlnAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGACCTTGATGATTTAGAAGAGCTTCTGTGAGAAAGATG 1639
DB 461 ValValSerIleProLysGlnThrTyrPyrAspLeuGlnGlnValLeuLeuGlnGlnMet 480
QY 1640 AAGATTTTCCGGAACCCAGCTGCTTATTCAGAAATGAGCTTCCACTAAGAGAGCAACA 1699
DB 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATGTTGTTCAACGGCTGGGGTGGCCAGCTCCCTTTCACCGGTGATATTTC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTTGGAGTGGTCT 1819
DB 521 GlyLysAlaCysAlaGlyCysCysLeuAlaArgAspProTyrCysAlaIleTrpAspGlySer 540
QY 1820 GCATGTTCCGCTATTTTCCCACTGCAAGAGAGCCCAAGAGCGCAAGATTAAGAAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysAspArgThrAlaArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCAGTCACTGATTCAGATCTTACACATGATTAATCACATGATGCCAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerThrPheuHisIleAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGAAATCATATGTTGTTAGAGAAATGATGACCAATTTTGAATGCGATCCGAG 1999
DB 581 GlnGlnArgIleIleTyrGlyValGlnAsnSerSerThrPheLeuGlnCysSerProLys 600
QY 2000 TGGCAGAGAGCGCTGGTCTATTTGGCAATTCAGAGAGCGCAATGAGAGCGAAAGAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnAspArgAspGlnArgLysGlnGln 620
QY 2060 ATCAGAGTGAATGATCATATATCATCAGAGCAGATCAAGGCTTCTGCTAGTGTACAA 2119
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuAspSerLeuGln 640
QY 2120 CAGAGAGATTCAGGCAATTAATCTGCGCATGGGTGAGAAATGGGTTCAATCAAACTCTT 2179
DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTTGGAGTCAATGACACAGAGCACTTTGGAAGAACTTTCATTAAGAT 2239
DB 661 LeuLysValThrLeuGlnValIleAspThrGlnHisLeuGlnGlnLeuLeuHisLysAsp 680
QY 2240 GATGATGAGAGATGGCTTAAGACCAAGAAATGTCCAATAGCATGACACTAGCCAGAG 2299
DB 681 AspAspGlyAspGlySerLysThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGACAGCTCATCAACCCCAATTCACACAGATGATGAG 2359

Db 701 ValTrrYrArGAspHneGlnLeuIleAsnHisProAsnLeuAsnThrIleAspGlu 720
 QY 2360 TTCTGTGAACAAGTTTGAAGAGGACCGGAAACAGCTGCGCAAGGCGAGACATACC 2419
 Db 721 PheCysGlnGlnValTrrYrAspAspArgLysGlnArgGlnArgProGlnHisIstr 740
 QY 2420 CCAAGGACAGTAAACAATGGAAGACCTTACAGAAATAAGAGTAAAGAGAGAG 2479
 Db 741 ProGlnYAsnSerAsnLysIrrYrShiIleuGlnGlnAsnLysGlnYrArgAsnArgArg 760
 QY 2480 ACCCAGCAATTTGAGAGGCGCACCCAGAGTGTG 2512
 Db 761 ThrHisGlnPheGlnArgAlaProArgSerVal 771
 RESULT 6
 US-09-791-537-31106
 ; Sequence 31106, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 31106
 ; LENGTH: 771
 ; TYPE: PRP
 ; ORGANISM: Homo sapiens
 US-09-791-537-31106

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-31106 (1-771)

QY 200 ATGGGCTGGTTAACTAGAGATGTCGTCTTTCTGGGAGTATTACTTACGAGAGAGA 259
 Db 1 MetGlyTrrPheThrArgLleValCysLeuPheTrrGlyValLeuThrAlaArgAla 20
 QY 260 AACTATCAGATGGAAGAACAATGCGCAAGGCTGAATTAATCTTACAAAGAAATGTTG 319
 Db 21 AsnYrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTrrLysGlnMetLeu 40
 QY 320 GAATCCAAATATGATGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATACCTTC 379
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTrrHisIstrPhe 60
 QY 380 CTTTGGATGAGAACGCGTAGGCTGTATGTTGAGCAAGAGATACATATTTTCATTC 439
 Db 61 LeuLeuAspGlnGlnArgSerArgLeuYrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACGTGTTAATATCAAGATTTTCAAAAGATTGTGTGCGCAGTATCTTACACAGAGAGA 499
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrrProValSerYrThrArgArg 100
 QY 500 GATGAATGCAAGTGGCTGGAAGAAAGACATCTCGAAGAAATGTGCTAATTTTCATCAAGTA 559
 Db 101 AspGlnCysLysIrrYrAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCAATATATAGACTCACTTGTAGCGCTGTGGAACGGGGCTTTTATCATCAAT 619
 Db 121 LeuLysAlaIrrYrAsnGlnThrHisLeuYrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTAATAATGGACATCATCTGAGAGCAATATTTTAACTGAGAGAACTCA 679

Db 141 CysThrYrLleGlnIleGlyHisHisProGlnAspAsnIlePheLysLeuGlnLysAsnSer 160
 QY 680 CATTTTGAAGACGGCGGGAAGAGTCCATATGACCTTAAGCTCTGACAGACATCCCTT 739
 Db 161 HisPheGlnAsnGlyArgGlyLysSerProYrAspProLysLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGATTAATTAATCTGGAATCTGAGCTGATTTTAATGAGGCGAGACTTTGCT 799
 Db 181 LeuIleAspGlyGlnLeuIrrYrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCGAACCTTTGGGACACCCACCAATACAGACAGACAGACATGATTTCCAGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGlnGlnHisAspSerArgTrp 220
 QY 860 CTCATATGATCCAAATTCATTAATGTCGCCACCTCATCTGAGAGCTGAGACATTCGAAGAT 919
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGlnSerAspAsnProGlnLysP 240
 QY 920 GACAAAGTAACTTTTCTCCGTGAAGAAATGCAATAGATGAGAGACACTCTGGAAGAGCT 979
 Db 241 AspLysValIrrYrPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGAAATAGTCAATATGCAAGAAATGACTTTGAGAGGCGACAGAAAGTCTGTTG 1039
 Db 261 ThrHisAlaArgLleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAATAGACAACTTCTCAAAGCTGCTGATTTGCTCAGTCCAGTCCAGTCCAAATGCG 1099
 Db 281 AsnLysIrrYrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTGATGAAGTGCAGAGATGATTCCTAATGAATTTTAAGATCCTTAA 1159
 Db 301 IleAspThrHisAspAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGATATGAGAGTGTTAACAATTCACAGTAAATTTTTCAGAGATCAGCCGCTG 1219
 Db 321 AsnProValValIrrYrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGATAGATAGATGATGAGAGAGGCTTCTGTCCTTGCCTATGACCAAGGAGAT 1279
 Db 341 CysMetYrSerMetSerAspAspValArgArgValPheLeuGlyProYrAlaHisArgAsp 360
 QY 1280 GGACCCCACTAATCAATGGGTGCTTATCAAGAAAGTCCCTATCCACGCGCAGAGAACT 1339
 Db 361 GlyProAsnYrGlnIrrYrValProYrGlnGlyArgValProYrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAACAATTTGGTGTGTTGACTCTACAAAGACCTTCAGATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGAAGAAGTCAATCCAGCCCATGTACAAATCCAGTGTTCCTATGAAACATGCCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetYrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGATGCAAAAGCGATGAATTAATATGAAATTTTACAAAATTTGTGTGACCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnYrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAGATGACAGATATGATATGTTATTCGGAACAGATGTTGGACCGTCTTAA 1579
 Db 441 AlaGlnAspGlyGlnIrrYrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTCAATTCCTAAGAGAATCTTGATGATTTTGAAGAGGCTTCTGCGAAGAAATG 1639
 Db 461 ValValSerIleProLysGlnThrTrrYrAspLeuGlnGlnValLeuLeuGlnGlnMet 480
 QY 1640 ACAGTTTTCGGAACGAGCTATTTTCAAGATGAGGCTTTCACATAGAGAGCAACA 1699
 Db 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATGGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGCTGTGATATTTAC 1759


```

Db      501 leutyrlleglserthrilaaglyalaaglnleuproleuhsargcysaspi1etyr 520
QY      1760 GGAAGACCGTGTGTGAGTGTGCTGCTGCGCCGAGACCTTACCTGTGTTGGATGTTCT 1819
Db      521 Gilyvalacysalaaglcyscyslevalaargasproryrcysalatpaspiglyser 540
QY      1820 GCATGTTCTGCTATTTTCCCACTGCAAGAAGACGCAAGACGCAAGATATTAAGAAAT 1879
Db      541 AlAcysaserargtrprrheprothrilaalyaharghrthrargharglnapilleahgasm 560
QY      1880 GGAAGACCGTGTGAGTGTGCTGCTGCGCCGAGACCTTACCTGTGTTGGATGTTCT 1939
Db      561 Gilyasprroleuhsrhlscysaserasprleuhsrhlscysaserasprleuhsrhlscysaser 580
QY      1940 GAAGAGATTCATCTTGTGTTAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1999
Db      581 Glnghlurgtlletlerygllyvalglnasrserthrphleuglucyssepprolys 600
QY      2000 TCCGACAGACCGGTGTCTATTTGGCAATTCAGAGCGGAATGAAGACGCAAGAAAGAGAG 2059
Db      601 Serghlnhrglaleuvaltyrrtprghlnphghlnhrglansnglunarglysglunghl 620
QY      2060 ATGAGAGTGAATGATCATATCATGACGACGATCAGAGCGCTTCTGCTAGCTTACTTACAA 2119
Db      621 lleargvalasprahsrllelleargthrAspglnghllyleuueuargserleughl 640
QY      2120 CAGAAGATTCAGAGATTAATCTCTGCGCATGCGGTGGAACATGGGTTTCAACAACTCTT 2179
Db      641 Glnlyasprserghlyantryleucyhsialaaglnlunlsglyprelleghlnthrleu 660
QY      2180 CTTAAGGTAACTCTGAGAGTCAATTCACACAGAGATTTGGAGAACTTCTTCAATAAGAT 2239
Db      661 leuhyvalthrleuughlvalilleasprthrghlhlleuughlueuhsrlybap 680
QY      2240 GATGATGAGATGAGCTTAAAGACCAAGAAATGTCATATGACATGACACTTACCGAGAG 2299
Db      681 Aspaspglyaspglyserlysthrlysglmeaserasrmetthrproserghllys 700
QY      2300 GTCTGTACAGAGACTTCATGACAGCTTCATCAACACCCCAATGTCAAACGATGATGAG 2359
Db      701 Valtrtyrtrargasprhmetghlnleuillleashnlsproasnlleuashntrmetasprghl 720
QY      2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGAAACAAAGTCGCGCAAGGCGCAGACATACC 2419
Db      721 PhcCysglunghlvaltrplysargasprarglysglhnhrghlghlnhrproghlysthr 740
QY      2420 CCGAGGAACGATTAACAATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAACAGAGAG 2479
Db      741 Proghlyasnsrseranlysthrlyshlsleughlghlunashlysglyargasnararg 760
QY      2480 ACCCAGAAATTTGAGAGGCGACCCAGAGAGTGC 2512
Db      761 Thrhsghlnphghlnarglalaarglalaarglalaarglalaarglalaarglalaargl 771

```

```

RESULT 7
US-10-067-632-54
; Sequence 54, Application US/10067632
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
;               Kolodkin, Alex L.
;               Matthes, David
;               Bentley, David R.
;               O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/067,632
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/835,268
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-067-632-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-067-632-54 (1-771)
QY      200 ATGGCGTGTGTTAATGAGATTTGTCTGTTTCTGCGGAGATTAATTAACAGCAAGACA 259
Db      1 MetghlytrpleuthrArghllevalCysleuphetrglyvalleuethrAlaargAla 20
QY      260 AACTTCGAAAGGGAAGAAACAATGTCGCAAGCGTGAAGATTAATCTTACAAAGAAATGTTG 319
Db      21 Aenlyrglnaenglylvalasnasnvalproahrgleuhsrlyseuethrlysglueu 40
QY      320 GAATCAACAATGTGATCATCTTCAATGAGCTTGGCAACAGCTCCAGTTATCATACCTTC 379
Db      41 GlnserasnasnvalillethrphasnglyleuAlaasnsrserthrhlsthrphe 60
QY      380 CTTTGTGATGAGAAACGAGTARGCTGTATGTTGAGCAAAAGATCAATATTTCAATTC 439
Db      61 leuueasprghlunghlarsgerargleutyrglyalalyasprhllepheserphe 80
QY      440 GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGCGCAAGTATCTTACACCGAGAGA 499
Db      81 Aspleuvalaenillelyasprhghlnlysllevaltrprovalserthrthrarghrg 100
QY      500 GATGATGACAGTGTGCGTGAAGAAAGACATCTGAAAGATGTGCTAATTTTCATCAAGTA 559
Db      101 AspglucylystprlaleglylsasprlleuuehsrlysglucyAlaasnrphelleysval 120
QY      560 CTTAAGCATATATATCAACTCATCTTGAACGCTGTGGAACGCGGCGCTTTTCATCAATT 619
Db      121 leuysalatyrgaenglnthrlthlsleutyrglAlaCysghlythrghlyalaphhehl 140
QY      620 TGCACCTCATTTGAATTTGACATCATCTGAGACATATTTTAACTGGAGAACTCA 679
Db      141 Cysthrtyrlleglnlleghlyhlsheprghlunspaanillephelyleughlunshrr 160
QY      680 CATTTGAAACGCGGTGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTT 739

```


CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 284
LENGTH: 771
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGCGTGTACTAGAGATGTCCTCTTCTGGGAGATTAATTACAGCAAGCA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuThrAlaArgAla 20
QY 260 AACTACAGATGAGAGAAACATATGACCAAGGCTGAATTAATCTACAAAGAAATGTTG 319
DB 21 AsnTrpGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTrpLysGlnLeu 40
QY 320 GAATCAACATGTGATCACTTCAATGCTTGCCCAACAGCTCAAGTTATCAATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTrpHisThrPhe 60
QY 380 CTTTGGATGAGAGACGAGATGCTGTATGTTGGACCAAGATCAATATTTTCAATTC 439
DB 61 LeuLeuAspGlnGluArgSerArgLeuTrpValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGATCTTAACACAGAA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValIleProValSerTrpHisArgAla 100
QY 500 GATGAATGCAAGTGGCTGAAAAAGACATCTGAAAAAGATGTCTAATTTTCAGAGTA 559
DB 101 AspGlnCysLysTrpIleArgLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCACTTGACGCTGTGGAAACGGGGGCTTTTCAATCA 619
DB 121 LeuLysAlaTrpAsnGlnThrHisIleLeuTrpAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGACCTAGTGAATTTGAGATCATCCCTGAGAGAAATATTTTAAGCTGAGAGACTGA 679
DB 141 CysTrpTrpIleGlnIleGlyHisIleProGlnAspAsnIlePheLysLeuGlnAsnSer 160
QY 680 CATTTTGAACCGCGCTGGAGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCT 739
DB 161 HisPheGlnAsnGlyArgGlyLysSerProLysAspProLysLeuThrAlaSerLeu 180
QY 740 TTAATATGATGAGATTAATTAATCTGGAATGCAAGCTGATTTTATGGGGGAGAGCTT 799
DB 181 LeuIleAspGlyGlnLeuTrpSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

QY 800 ATCTCCGAACCTTTGGGACCAACCACCAATACAGAGACAGACATGATTTCCAGTTGG 859
DB 201 IlePheArgThrLeuGlyHisHisIleProIleArgTrpGlnHisIleAspSerArgTrp 220
QY 860 CTCATATGATCAAAAGTCTATTAGTCCCACTCATCTCAGAGAGTGAACATCTTAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisIleuIleSerGlnSerAspAsnProGlnAsp 240
QY 920 GACAAAGTATTAATTTTCTCCGTAAGAAATGAATGAGACACTTCGGAAAGCT 979
DB 241 AspLysValTrpPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAAGCTAGAAATAGTCAATATGCAAGATGACTTTGGAGGCAAGAACTGGTG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyHisIleArgSerLeuVal 280
QY 1040 AATAAATGACAACAATCTCTCAAAAGCTGCTGATTTGCTCAGTCCAGGTCCAATGGC 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGAATGAATGCAAGATGATTCCTTAATGAATTTTAAGATCTTAA 1159
DB 301 IleAspThrHisPheAspGlnLeuGlnAspValIlePheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGATATGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGATCAAGCGTG 1219
DB 321 AsnProValValTrpGlyValAlaPheThrTrpSerSerAsnIlePheLysGlySerIleVal 340
QY 1220 TGTATGATAGCATGATGATGTGAGAAAGGTTGCTTGGTCCATATGCCCAAGGAT 1279
DB 341 CysMetTrpSerMetSerAspValArgArgValIlePheLeuGlyProTrpAlaHisArgAsp 360
QY 1280 GGACCCCACTAATCATATGGTGGCTTATCAAGAAAGCCCTATCCAGGCAAGACT 1339
DB 361 GlyProAsnTrpGlnTrpValProTrpGlnGlyArgValProTrpProArgProGlyThr 380
QY 1340 TGTCCAGCAAAAACATTTGGTGTGTTTGAATCTTCAAAAGAACCTTCGATGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTGCAAGAGATCATCCACCCATGTAACAATCCAGTGTTCCTATGAACAATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTrpAsnProValIlePheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGAGATGTAATTAATTAACAATTCGTGTGACCGAGATGAT 1519
DB 421 IleValIleLysThrAspValAsnTrpGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGTATGATGTTATGTTATCGGAACAGATGTGGACGCTTTTAAA 1579
DB 441 AlaGlnAspGlyGlnTrpAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTCAATTCCTAAGAGACCTTGATGATTTTGAAGAGGTTCTGCTGAGAGAAATG 1639
DB 461 ValIleSerIleProLysGlnThrTrpTrpAspLeuGlnGlnValLeuLeuGlnGlnMet 480
QY 1640 ACAATTTTTCGGGAACGACATGCTATTTTCAGCAATGAGAGCTTTCACATGAAGAAAGAA 1699
DB 481 ThrValIleAspArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGln 500
QY 1700 CTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
DB 501 LeuTrpIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleLys 520
QY 1760 GGGAAAGCGTGTGCTAGAGTGTGCTGCGCCGAGACCTTAATCTGTGCTGGAGTGTCT 1819
DB 521 GlyLysAlaCysAlaGlnCysCysLeuAlaArgAspProTrpCysAlaIleTrpAspIleLys 540
QY 1820 GATGTTCTCGTATTTTCCACTGCAAGAAAGACGCAACAAGCAAGATTTAAGAAAT 1879
DB 541 AlaCysSerArgTrpPheProThrAlaLysArgArgTrpArgArgGlnAspIleArgAsn 560

```

QY 1880 GGAGACCCACTGACTACTGTTTACACTTAACCATGATATACCATGCGCACACCTT 1939
    |||
Db 561 GtYasPProleuthrthiScySseSerAspLeuNh1sh1AspAsnh1sh1sglYh1SsePro 580
    |||
QY 1940 GAAGAGAGATCATCTATGTTGTAGAGATAGAGACATTTTGTGAATAGCGCTCCGAG 1999
    |||
Db 581 GtUGluArg1Le1LeYrGLyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
    |||
QY 2000 TCGCAGAGACCGCTGCTTATTTGGCAATTCAGAGCGCAATGAAGAGCAAAAGAGAG 2059
    |||
Db 601 SerGlnArgAlaLeuValYrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
    |||
QY 2060 ATCAGAGTCGATGATCATATCATCAGACAGATCAAGCGCTTCTGCTACGATGATCA 2119
    |||
Db 621 IleArgValAspAspsh1s1le1le1ArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
    |||
QY 2120 CAGAAGGATTCAGCAATTAACCTTCGCCCATGGCGTGGAAATGGGTTCTATCAAACTTT 2179
    |||
Db 641 GlnLysAspSerGlyAsnTyLeuCysh1s1AlaValGlnH1sGlyPhe1leGlnThrLeu 660
    |||
QY 2180 CTTAAGGTAAACCTGGAGATGCTATTCACAGAGCATTTTGAAGAACTTCTTATCAAAAGAT 2239
    |||
Db 661 LeuLysValThrLeuGlnValIleAspThrGluH1sLeuGlnGluLeuNh1sLysAsp 680
    |||
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGCATGACACTTACCGCAGAG 2299
    |||
Db 681 AspAspGlyAspGlySerLysThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
    |||
QY 2300 GTCTGTGACAGAGCTTCTATGAGCTCATCAACCCCAATCTCAACGATGATGAG 2359
    |||
Db 701 ValTrpTyArgAspPheMetGlnLeuIleAsnH1sProAsnLeuAsnThrMetAspGln 720
    |||
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGAGCGAAACCAACGCGCAAGCGCAGACATACC 2419
    |||
Db 721 PheCySGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlnH1sThr 740
    |||
QY 2420 CCAGGAAACAGTAAACAATGAGAGCACTTACAAGAAATTAAGAAAGTAAAGAGAGAG 2479
    |||
Db 741 ProGlnLysSerAsnLysTrpLysH1sLeuGlnGlnLysAsnLysGlyArgAsnArgArg 760
    |||
QY 2480 ACCCAAGCAATTTGAAGGGCACCAGAGAGTGC 2512
    |||
Db 761 ThrH1sGlnPheGlnArgAlaProArgSerVal 771
    |||
RESULT 9
US-10-247-671-164
; Sequence 164, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 164
; LENGTH: 771
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1930967CD1
US-10-247-671-164
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0

```

```

Beet Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 28 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-247-671-164 (1-771)
QY 200 ATGGCGTGGTTAACTAGAGATTGCTGTCTTTTCTGGGAGATTAATTACAGCAAGACA 259
    |||
Db 1 MetGlyTrpLeuThrArg1LeValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
    |||
QY 260 AACTTCAGATGGGAAAGAACATGTCGACAGCTGAAATTTCTTACAAAGAAATGTTG 319
    |||
Db 21 AsnTyrgLysAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyLysGlnMetLeu 40
    |||
QY 320 GAATCAACAATATGATACCTTCAATGCGCTGGCCACAGCTCCAGTATCATCCTTC 379
    |||
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyH1sThrPhe 60
    |||
QY 380 CTTTTGATGAGAACGAGATGAGCTGTATGTTGGAGCAAGATCACATATTTTCATTC 439
    |||
Db 61 LeuLeuAspGluGluArgSerArgLeuTyArgAlaLysAspH1sIlePheSerPhe 80
    |||
QY 440 GACCTGTTAATATACAGATTTTCAAAAGATGTGCGCAGATCTTACACCAAGAGA 499
    |||
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyThrArgArg 100
    |||
QY 500 GATGAATCAAGTGGCGCTGAAAAGACATCCGAAAGAAATGCTAATTTGATCAAGGTA 559
    |||
Db 101 AspGlnCysLysTrpAlaGlyLysAspLleLeuLysGlnCysAlaAsnPhe1leLysVal 120
    |||
QY 560 CTTAAGGCATTAATACAGACTCACTTGTACCGCTGTGAGACGGGGCTTTTCATCCAAAT 619
    |||
Db 121 LeuLysAlaTyArgAsnGlnThrH1sLeuTyArgLysGlyThrGlyAlaPheH1sProIle 140
    |||
QY 620 TGCACATCAATTGAATGACATCATCTTAAGACCAATATTTTAAGCTGAGAACTCA 679
    |||
Db 141 CysThrTyTrl1leGln1leGlyH1sH1sProGlnAspAsnIlePheLysLeuGlnAsnSer 160
    |||
QY 680 CATTTTGAACCGGCGCGGAGAGAGCTCAATATGACCTTAAGCTGACAGCATCCCTT 739
    |||
Db 161 H1sPheGluAsnGlyArgGlyLysSerProTyArgProLysLeuLeuThrAlaSerLeu 180
    |||
QY 740 TTAATAGATGAGAAATTAATCTCTGAACTGCACTGATTTTATGGGCGAGACTTGTCT 799
    |||
Db 181 LeuIleAspGlyGluLeuTySerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
    |||
QY 800 ATCTTCGCACTCTTGGGACACCAACCAATACAGACAGACAGCATGATTCAGGTGG 859
    |||
Db 201 IlePheArgThrLeuGlnH1sH1sH1sProIleArgThrGlnH1sAspSerArgTrp 220
    |||
QY 860 CTCATGATGCAAGTTCATTAAGTGGCCCACTCATTCAGAGAGTGAACATCCGAGAGAT 919
    |||
Db 221 LeuAsnAspProLysPheIleSerAlaH1sLeuIleSerGlnUserAspAspProGlnAsp 240
    |||
QY 920 GACAAAGTATACCTTTTCTCCGTGAATAATCAATAGATGAGAGAACTCTGAAAAAGCT 979
    |||
Db 241 AspLysValTyrrPhePhePheArgGluAsnAlaIleAspGlyGlnH1sSerGlyLysAla 260
    |||
QY 980 ACTCAGCTAGATAGTGCATATGCAAGATGACATTTTGAAGGACAGAAAGTCTGGTG 1039
    |||
Db 261 ThrH1sAlaArg1leGlyGln1leCysLysAsnAspPheGlyGlnH1sArgSerLeuVal 280
    |||
QY 1040 AATTAATGACAAACATTCCTTAAGCTCGTGTATTTGCTCAGGCGCAGGCTCCAAATGGC 1099
    |||
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
    |||
QY 1100 ATTGACACTCATTTTGAATGAACTCAGAGATATATCTCAATGAAGACTTAAAGATCTAAA 1159
    |||
Db 301 IleAspThrH1sPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
    |||
QY 1160 AATCAGGTGTATATGAGATGTTTACGATTCACAGTAAATTTTCAAGGATCAGCGGTG 1219
    |||
Db 321 AsnProValValTyrgLysValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
    |||

```

```

QY 1220 TGTATGATAGCATGAGTGTGAGAAAGGTTCTTGATCCATATGCCCAGAGGAT 1279
    |||
Db 341 CymetylserinetserrapvalArgArgValPheuleuLProtyrAlaHisArgAsp 360
QY 1280 GGAACCAATCATCAATGGTGGCTTATCAAGAAAGTCCCTATCCAGGCCAGAACT 1339
    |||
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyTyr 380
QY 1340 TGTCCAGCAAAACATTGGTGGTTTGAAGCTTCAAGAAAGCCTTCTGATGTATTA 1399
    |||
Db 381 CysProSerTyrThrPheGlyGlyPheAspSerThrTyrAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGATCATCCAGCATGTATCAATCCAGTGTCTTGAACAATCCGCA 1459
    |||
Db 401 ThrPheAlaIArgSerHisProAlaMetTyrAsnProValPheProMetAsnAlaArgPro 420
QY 1460 ATATGATCAAAAACGATGTAAATTTCAATTTTCAACAAATTGTCTGATGCCAGTGGAT 1519
    |||
Db 421 IleValIleTyrThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGATGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
    |||
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGAGACTGGATGATTTAGAAAGGTTCTGCTGAGAAAGT 1639
    |||
Db 461 ValValSerIleProLysGlnTyrTrpTyrAspLeuGlnValIleuLeuGlnLumet 480
QY 1640 ACAGTTTTTGGGAACCGACTGCTATTTACAGCAATGAGCTTTTCACTAAGACAA 1699
    |||
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrTyrGlnGlnGln 500
QY 1700 CTATATATGTTGTTCAACGGCTGGGTTGCCAGCTCCCTTACACGGTGTATTTAC 1759
    |||
Db 501 LeuTyrTleIleYserThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTCTTGGATGGTCT 1819
    |||
Db 521 GlyLysAlaCysAlaGluCysGlyLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAGACAGCAAGATATTAAGAAAT 1879
    |||
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgTyrThrArgGlnAspIleArgAsn 560
QY 1880 GGAAGACCACTGACTACTGTTGAGCTTACACCATGATTAACATGAGCCACGCTT 1939
    |||
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGAAATAGTAGACATTTTGGAAATGCAAGTCCGAG 1999
    |||
Db 581 GlnGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAGAGCGAAAGAAAGAG 2059
    |||
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGlnGluArgLysGlnGlu 620
QY 2060 ATCAGAGTGATATATATATCATCAGAGACATCAAGCGCTTCTGCTAGTATGATCAAA 2119
    |||
Db 621 IleArgValAspAspHisIleIleArgTrnAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAAGATTCAGGCAATTAACCTCTGCAATGCGGTGGAACATGGTTCATCAAACTCTT 2179
    |||
Db 641 GlnLysAspSerClyAsnTyrLeuCyHisAlaValGlnHisGlyPheIleGlnTrnLeu 660
QY 2180 CTTAAGATTAACCTGGAAGTCAATGACAGAGCAATTTGGAAGAACTTCTTCAATAAGAT 2239
    |||
Db 661 LeuLysValIleThrLeuGlnValIleAspTrnGlnHisIleuGlnGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGAGATGCGCTTAAGACCAAGAAATGCTCAATAGCATGACACTTAGCCGAGAG 2299
    |||
Db 681 AspAspGlyAspDlySerLysThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700

```

```

QY 2300 GTCTGTCACAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGAG 2359
    |||
Db 701 ValIleTyrArgAspPheMetClnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGAGCCGAAAACAAGTCCGCAAGGCGACAGATACC 2419
    |||
Db 721 PheCysGlnGlnValIleTyrPheArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAAACAAATGGAAGCACTTACAGAAAATTAAGAAAGTGAACAGAGAG 2479
    |||
Db 741 ProGlyAsnSerAsnLysTrpLysHisIleuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAAGATTTGAGAGGCAACCCAGAGTGTCT 2512
    |||
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

```

```

RESULT 10
US-10-262-538-10
; Sequence 10, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alltalo et al
; TITLE OF INVENTION: NEUROBILIN/VEGF-C/VEGR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28367/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

```

```

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 28 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)

```

QY 200 ATGGCTGCTGTTAATCAGATGATGCTGTCTTTCTGGGAGTATTAATCAAGACAGCA 259
    |||
Db 1 MetGlyTrpLeuThrArgIleValCysLeuThrTrpGlyValLeuLeuThrHisArgAla 20
QY 260 AACATACGAATGGGAAGAACATGTGCCAAGGCTGAATATTCCTACAAAGAAATGTTG 319
    |||
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGATGATCACTTCAATGGCTTGGCCAAACAGCTCAGTATCATACCTTC 379
    |||
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATTC 439
    |||
Db 61 LeuLeuAspGlnGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAATATCAAGATTTTCAAAAAGATTGTGGCCAGTATCTTAACACAGAGA 499
    |||
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 CATGAATGCAAGTGGGCTGGAAGAAACATCTGTAAGAAATGTGCTAATTTCAATCAAGTA 559
    |||
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGCATTAATCAAGACTCACTTGAACGCTGTGGAAGAGGGGGCTTTTCATCAATT 619
    |||
Db 121 LeuLysAlaTyrAsnGlnThrHisIleuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAACTGAGAGAACTCA 679
    |||

```

Db 141 CysThrTyrIleGluIleGlyHisIleProGluAspAsnIlePheIleuGluAsnSer 160
Qy 680 CATTTTGAAGAGCGCGTGGAGAGAGTCCATATGACCTTAAGCTGTGACAGACCTCTT 739
Db 161 HisPheGluAsnGlyArgGlyIleSerProTyrAspProIleuLeuThrAlaSerLeu 180
Qy 740 TTAATATGATGAGATTTATATCTCTGGAACGTGACGTGATTTATGGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCGGAACCTTTGGGACACCAACCAATCGAGACAGAGATGATATTCAGAGTGG 859
Db 201 IlePheArgThrLeuGlyHisIleHisIleProIleArgThrGluGlnHisAspSerAlaGlyTyr 220
Qy 860 CTCATATGATCCAAAGTTCATTAAGTCCCACTCATCTCAAGAGTGAACATCCTGAAGAT 919
Db 221 LeuAsnAspProIlePheIleSerAlaHisIleuIleSerGluSerAspAsnProGluAsp 240
Qy 920 GACAAAGTATACCTTTTCTTCGCTGAAAATGCAATATGATGAGAAACACTCTGGAAAAGCT 979
Db 241 AspIleValTyrPhePhePheArgIleAsnAlaIleAspGlyGluHisSerGlyIleAla 260
Qy 980 ACTTCACGCTAGAAATAGTATGATATGCAAGATGATGATCTTGGAGGCGACGAAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysIleAsnAspPheGlyGlyHisArgSerLeuVal 280
Qy 1040 AATAAATGACAAACATTCCTCAAGCTCGTCTGATTTGCTGAGTGCAGTCCAAATGGC 1099
Db 281 AsnIleGlyThrThrPheLeuIleAlaGluIleCysSerValProGlyProAsnGly 300
Qy 1100 ATGACACTCATTTTGTGATGACAGTGCAGATGATATCTTAATGAACTTTAAGATCCTTAA 1159
Db 301 IleAspThrHisPhePheGlyIleuGlnAspValPheLeuMetAsnPheIleAspProIle 320
Qy 1160 AATCCAGTTGTATATGAGTGTATACGACTTCCAGTAAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyAlaPheThrThrSerSerAsnIlePheIleGlySerAlaVal 340
Qy 1220 TGTATATAGCATGATGATGAGAGAGGTGTCTTGTCTGATGATCCCAAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValAlaGargValPheLeuGlyProTyrAlaHisArgAsp 360
Qy 1280 GGAACCAATCATCATGAGGTGCTTATCAAGAAAGATCCCTATCCAGGCGCAGAGACT 1339
Db 361 GlyProAsnTyrGlnThrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
Qy 1340 TGTCCACGAAACATTTGGTGTGTTTGAAGTCTTCAAGAGACCTTCTGATATGTTTATA 1399
Db 381 CysProSerIleThrPheGlyGlyPheAspSerThrIleAspLeuProAspAspValIle 400
Qy 1400 ACCCTTGAAGAAGTATCAGCATGATGATCCAGTGTCTTCTTAAGAAATGCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy 1460 ATATGATCAAAAACGAGTAAATTAATTAACAATTTGCTAGAACGAGTGCAT 1519
Db 421 IleValIleIleThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy 1520 CGAAGAAGTGAACAGTATGATGTTATGTTATCGAAGACAGATGTTGGACGCTTTAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuIle 460
Qy 1580 GATAGTTCAATTCCTTAAGAGACTGCTATGATTTTGAAGAGTCTGCTGGAAGAAAG 1639
Db 461 ValIleValSerIleProIleGlyIleThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
Qy 1640 ACAAGTTTTCGGAACCGACGCTATATTCAGCAATGAGACTTTCACATAAGACGACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrIleGlnGlnGln 500
Qy 1700 CATATATTTGTTCAACGCGTGGGTTGCCAGCTCCCTTTACACCGGCTGTGATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyAlaIleGlnLeuProLeuHisArgCysAspIleTyr 520

Qy 1760 GGAAGAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCTTAAGTGTGGATGTTCT 1819
Db 521 GlyIleValCysAlaGlyCysIleuAlaArgAspProTyrCysAlaThrAspGlySer 540
Qy 1820 GCATGTTCTCGCTATTTTCCCATCTGCAAGAGACGCAACAGACGACATTAAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaIleValArgAlaGlnThrArgGlnAspIleArgAsn 560
Qy 1880 GGAGACCACTGACCTGCTGACATTCACCTTACCCATGATTAACACAGCCGACGACCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisIleAspAsnHisIleGlyHisSerPro 580
Qy 1940 GAAGAGACATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1999
Db 581 GlnGluValGlyIleIleTyrGlyValGlnIleAsnSerSerThrPheLeuGluCysSerPro 600
Qy 2000 TCGCAGAGACGCTGTCTTATTTGCAATTCAGAGCGCAATGAAAGCGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAlaGlnGluGluArgGlyGln 620
Qy 2060 ATCAGAGTGAATGATCATATCATCATGACAGATGACAGGCTTCTGCTACGTATCA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuAspSerLeuGln 640
Qy 2120 CAGAAGATTCAGCAATTAACCTCTGACATGCGGTGGAACATGGGTTCATACAACTCT 2179
Db 641 GlnIleAspSerGlyAsnTyrIleuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
Qy 2180 CTTAAGGTAACTTGAAGTCAATTGACAAGACATTTGGAAGAACTTCTCATTAAGAT 2239
Db 661 LeuIleValThrLeuGluValIleAspThrGlnHisIleuGlnGluLeuLeuHisIleAsp 680
Qy 2240 GATGATGAGATGCTCTTCAAGACCAAGAAATGTCTCATATGATACACTTACCGCAGAG 2299
Db 681 AspAspGlyAspGlySerIleThrIleGlnIleuSerSerAsnMetThrProSerGlnIle 700
Qy 2300 GTCTGTACAGAGACTTATGATGAGTGCATCAACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnIleuIleAsnHisProAsnLeuAsnThrMetAspGln 720
Qy 2360 TTCTGTGAACAAGTTTGAAGAAAGGAGCCGAAACACAGCTCGCAAGGCGCAGACATACC 2419
Db 721 PheCysGlnGlnValTrpIleArgAspArgIleGlnArgGlnArgProGlyHisThr 740
Qy 2420 CCAAGGAACGATTAACAATGGAAGCACTTACAGAAATTAAGAAAGTGAAGACGAGG 2479
Db 741 ProGlyAsnSerAsnIleTrpIleHisIleuGlnGluAsnIleGlyArgAsnArgArg 760
Qy 2480 ACCCAGCAATTTGAGAGGCAACCCGAGAGTGC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 11
US-60-323-784-164
; Sequence 164, Application US/60323784
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiftman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 P
; CURRENT APPLICATION NUMBER: US/60/323,784
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO: 164
; LENGTH: 771
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

OTHER INFORMATION: Incycle ID No: 1930967CD1
US-60-323-784-164

Alignment Scores:

Pred. No.:	0	Length:	771
Score:	771.00	Matches:	771
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	86.15%	Indels:	0
DB:	31	Gaps:	0

US-09-774-490-1 (1-2709) x US-60-323-784-164 (1-771)

```

QY 200 ATGGGCTGTTAACTAGATTTGCTGCTTTTCTGGGAGATTAATTACTTACAGCAAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGAGAAACAATGTCACAGCTGAATTAATCTACAAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysSerTyrLysGlnMetLeu 40
QY 320 GAATCCACAATGATGACATTTTCATAGGCTTGCCCAAGCTCCACTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrTrpAlaSerPhe 60
QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAGGATCAGATATTTCATTC 439
Db 61 LeuLeuAspGlnGluLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATTTACCCAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrTrpArgArg 100
QY 500 GATGATGCAAGTGGCTGGAAGAAAGACATCTGAAAGATGTGATATTCATCAAGTA 559
Db 101 AspGlnLysSerTrpAlaGlyLysAspIleLeuLysGlyCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGATATATATCAGACTCTTGTACGCTGTGGAACGGGGCTTTTCACTCAAT 619
Db 121 LeuLysAlaLysTrpAsnGlnThrHisLeuTyrAlaCysGlyLysTrpAlaPheHisProIle 140
QY 620 TGCACTACATGAAATGATGACATCATCTGAGAGCAAAATTTTAACTGGAGAACTCA 679
Db 141 CysThrTyrTrpIleGlnIleGlyHisIleAspGlnAspAsnIlePheLysLeuGlnAsnSer 160
QY 680 CATTTTGAAGACGGCGCTGGGAGAGTCCATATGACCTTAAGTCTGAGAGACATCCCT 739
Db 161 HisPheGlnAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATTAATTACTCTGGAATGCACTGATTTTATGGGGCAGACTTGGCT 799
Db 181 LeuIleAspGlyGlnLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATTTTCCGAATCTTTGGGACCAACCAACCAATGACAGACAGACAGCATGATTCAGAGTGG 859
Db 201 IlePheArgThrLeuGlyHisIleHisIleProIleArgThrGlnGlnHisIleAspSerArgTrp 220
QY 860 CTCATGATGCAAGATTCATTAATGATGAGTCCATCATCTGAGAGTGAATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisIleLeuIleSerGlnSerAspAsnProGlnAsp 240
QY 920 GACAAAGTACTTTTCTTCCGTGAATGCAATAGATGAGAACTCTGGAAGAACT 979
Db 241 AspLysValTyrTrpPhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATGAGTCAATATGCAAGATGACTTTGGAGGCAACAGAACTGTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGCAACAATTTCTCAAGCTCTGTGATTTGCTCAGTGGCAAGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

```

```

QY 1100 ATTGACACTCAATTTGATGAGTCAAGATGATTAATCTTAATGAACTTTAAAGTCCATAA 1159
Db 301 IleAspThrHisPheAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTTTACAGCTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyAlaPheThrTrpSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATAGATGATGATGATGAGAGAGGTTGCTTGGTGCATATGATCCACAGGGAT 1279
Db 341 CysMetTrpSerMetSerAspValAspArgValPheLeuGlyProTyrAlaHisAspAsp 360
QY 1280 GACCCCACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGGCGCAGAGACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGTGTTGACTCTTACAAAGGACCTTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerTrpLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGAAGAGTATCCAGCATGTACATCCAGTGTTCCTATGAAACAATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnPro 420
QY 1460 ATAGTATCAAAACGAGATGTAATTAATTAATTAACAATTTGCTAGACCGAGTGAAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGTATGATGTTATGTTATCCGAAACAGATGTTGGACCGTCTTAA 1579
Db 441 AlaGlnAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGACTGTGATGATTTAGAAGAGCTTCGTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGlnThrTrpTyrAspLeuGlnGlnValLeuLeuGlnGlnMet 480
QY 1640 ACAATTTTTCGGAACCGACTGCTATTTCAAGAAAGAGCTTTCACATTAAGCAGACA 1699
Db 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTATATTTC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyAlaAlaGlnLeuProLeuHisIleArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGAGAGTGTGCTCCGCGCGAGACCTTATGTGCTTGAGAGTGTCT 1819
Db 521 GlyLysAlaCysAlaGlnCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAACAAGACGACAAGATTAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgAlaGlnThrArgGlnAspIleArgAsn 560
QY 1880 GGAAGCCCACTGACTCACTGTTCAACTTACACCAATGATATACCAATGGCCACAGCCT 1939
Db 561 GlyAspProLeuThrHisIleCysSerAspLeuHisIleAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGAAATCATATATGCTGTAGAGATATAGACATTTTGAATGAGAGTCCGAAG 1999
Db 581 GlnGlnLysGlyIleIleTyrGlyAlaGlnAsnSerSerTrpPheLeuGlnCysSerProLys 600
QY 2000 TCGCAGAGACGCTGGTCTATTGGCAATTCAGAGCGCAATGAAGAGGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnAlaGlnArgAsnGlnGlnArgLysGln 620
QY 2060 ATCAGAGTGGATGATCATATCATCAAGACAGATCAAGGCTTCTGTACGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuAspSerLeuGln 640
QY 2120 CAGAAGATTCAGGGAATTAATCTGCAATGGCGGAGCAACATGGGTCTATCAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCGAAGTCAATGACACAGAGCAATTTGGAAGAACTTTCATTAAGAT 2239

```



```

Db      661  LeuLVsValThLeuGLuValIleAspThrGluHsLeuGLuGluLeuHsIlySAsp 680
QY      2240  GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACCTTACGCCAGAG 2299
Db      681  AspaSpGlyAspGlySerIlySerIlySyluMeSerAspSerMetThrProSerGlnIlyS 700
QY      2300  GTCTGTACAGAGACTTCATGAGACTCATCAACCAACCCCAATCTCAACAGATGATGAG 2359
Db      701  ValTPlyrArGAspPheMetGlnLeuIleAsnHsIProAsnLeuHsThrMetAspGlu 720
QY      2360  TTCTGTGAACAAGTTTGGAAAAGGACCAAAACAACTGTGGCAAAAGCCAGACATACC 2419
Db      721  PheCySgluGlnValTPlyrAspArgSyluGlnArgGlnArgProGlyHsIlyThr 740
QY      2420  CGAGGAACAGTAAACAATGGAAGCACTTACAAAGAAATTAAGAGCTGAAGAACGAGAG 2479
Db      741  ProGlyAspSerAspIlySerIlySyluMeSerAspSerMetThrProSerGlnIlyS 760
QY      2480  ACCCAGCAATTTGAGAGGCGACCCAGAGAGTGC 2512
Db      761  ThrHsGluPheGluArgIleProArgSerVal 771

RESULT 12
US-09-791-537-32993
; Sequence 32993, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Dancet, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 32993
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-32993

```

```

Alignment Scores:
Pred. No.: 2,746-280 Length: 287
Score: 287.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.07% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-32993 (1-287)

```

```

QY      1652  GAACGAGCTGCTATTTCAGCAATGAGCTTTCCACTAAGCAGCAACAATATATATGGT 1711
Db      1  GluProThrAlaIleSerIleMetGlnLeuSerThrIlySgluGlnGlnLeuIlyrIleGly 20
QY      1712  TCAAGCGCTGGGCTGCCAGCTCCCTTTACACCGGTGTGATTTTACGGGAAAGCGTGT 1771
Db      21  SerThrAlaGlyValAlaGlnLeuProLeuHsIlySyluMeSerAspIleIlyrIlySAlaCyS 40
QY      1772  GGTGAGTGTGGCTGCGCCGAGACCTTACTGCTGGTGGAGTGGTGTTCGATGTTCTGCC 1831
Db      41  AlaGluCySyluLeuAlaArgAspProIlyrCySAlaTPAspGlySerIleAcylSerArg 60
QY      1832  TATTTTCCCACTGCAAGAGACGCAACAAGACGACAAGATATTAAGAAATGAGACCCACTG 1891
Db      61  TyrPheProThrAlaIlySArgIlyrThrArgArgGlnAspIleArgAsnIlyAspProLeu 80
QY      1892  ACTCATGTTTCACTTACACATGATATATCAACATGGCCAGACAGCCCTGAAGAGAAATC 1951
Db      81  ThrHsCySAspSerAspLeuHsIlySAspAsnHsIlySgluHsIleSerProGluGlnArgIle 100

```

```

QY      1952  ATCTATGCTGTAGAGAAATAGTAGCACTATTTTGAATGACAGTCCGAGAGTCCGAGAGCG 2011
Db      101  IleTyrGlyValGluAsnSerSerThrPheLeuGlnIlyCySerProIlySerGlnArgIle 120
QY      2012  CTGGCTATTTGGCAATTTCCAGAGCGGAATGAAAGCCGAAAAAGAAAGATCATGATGGAT 2071
Db      121  LeuValIlyrTPGlnPheGlnArgIlyrGlnArgAsnGlnIlyrGlySgluGlnIleArgValAsp 140
QY      2072  GATCATATCATCAGAGACAGATCAAGACCTTCTGCTACGAGTCAACAGAAAGATTCA 2131
Db      141  AspHsIleIleArgThrAspGlnIlyrLeuLeuMetAspIlyrSerLeuGlnIlySAspSer 160
QY      2132  GGCATTAACCTCTGCGCATCGCGTGGACATGAGTTCATACAACTCTTTAAGGTAAAC 2191
Db      161  GlyAsnIlyrLeuCyHsIleAlaValGlnIlySgluPheIleGlnThrLeuLeuValThr 180
QY      2192  CTGGAAGTCTTGAACACAGAGCATTTGGAAGAACTTTCTCATTAAGATGATGAGAT 2251
Db      181  LeuGluValIleAspThrGlnHsLeuGlnIlyrLeuLeuHsIlySAspAspAspGlyAsp 200
QY      2252  GGCCTAAGACCAAGAAATGTCCATAGCAATGACATGACATGACAGAGTCTGTGATACA 2311
Db      201  GlySerIlyrIlyrIlySyluMeSerAspSerMetThrProSerGlnIlyrValTPlyrArg 220
QY      2312  GACTTCATGACAGCTCATGACCAACCCCAATCTCAACAGATGATGATGATGATGATGAT 2371
Db      221  AspPheMetGlnLeuIleAsnHsIProAsnLeuHsThrMetAspGluPheCySgluGln 240
QY      2372  GTTTGGAAAAGGAGACCGAAACACAGTGGCAAAAGGCCAGACATATCCCGAGGAACGT 2431
Db      241  ValTPlyrAspArgSyluGlnArgIlyrGlnArgIlyrGlnArgProGlyHsIlyrProGlyAsnSer 260
QY      2432  AACCAATGAGACACTTCAAGAAATTAAGAAATGAGTGAAGAACAGAGAGACCAAGATTT 2491
Db      261  AsnIlyrTPlyrHsIlyrGlnGlnIlyrAsnIlyrSyluGlyArgAsnArgIlyrThrHsIlyrPhe 280
QY      2492  GAGAGGCGACCCAGAGAGTGC 2512
Db      281  GluArgIleProArgSerVal 287

```

```

RESULT 13
US-09-724-676-66689
; Sequence 66689, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 66689
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66689

```

```

Alignment Scores:
Pred. No.: 4,456-263 Length: 272
Score: 270.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.17% Indels: 0
DB: 21 Gaps: 0

```

```

US-09-774-490-1 (1-2709) x US-09-724-676-66689 (1-272)
QY      200  ATGGGCTGGTTAACTAGAGATGTCTGTTCTTGGGAGTATTACTTACAGACAGAGCA 259
Db      1  MetGlyrTPleuThrArgIleValCySleuPheThrProGlyValIleuLeuThrAlaArgAla 20
QY      260  AACTATCAGAAATGGAGAACAAATGTGCCAAGGCTGAATTAATCTTACAAAGAAATGTTG 319

```



```

Db      21  AenlyrglnasngilylsaaasnaValProArgluLeuylseuSerTyrlysglumeLeu 40
QY      320  GAATCCAACAATGATGATCACTTTCATATGCTGGCCCAACAGCTCCAGTTTCATACCTTC 379
Db      41  GluseAsnaValIlethrphenngilyleuAlaasnserserSerTyrhisrhphe 60
QY      380  CTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAATATTTTCATTC 439
Db      61  LeuleuAspGluGluarGserarGleuTyrValGlyAlaYsaPhisrIlepheserPhe 80
QY      440  GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATTTACCAAGAGA 499
Db      81  AspleuValAsnIleuYsaPheGlnlyValIleValIleProValserTyrThrArgArg 100
QY      500  GATGAATGCAAGTGGGCTGGAAGAAGACATCTGAAAGATGTGCTAATTTTCAACAGTA 559
Db      101  AspgIuCylystrpAlagIlysaPrlleuYsaIleuYsaIleuYsaIleuYsaIleuYsa 120
QY      560  CTTAAGGCATATATACAGACTCACTGTACGCTGTGGAACGAGGCTTTTCATCCAAAT 619
Db      121  LeuYsaIaIaYsaGlnThrhisIleuYsaIaIaCysglYThrGlyAlaPhehisProIle 140
QY      620  TGCACTCACTAATTAATGAGCATCATCTGAGAGCAATTTTAACTGGAGAACTCA 679
Db      141  CysThrTyrIleGlnIleGlyhisIshIspIroGluAspAsnIlePheYsaIleuGluAsn 160
QY      680  CATTTGAAAACGCGCTGGGAAGAGTCATATGACCTTAAGCTGAGCAGCATCCCTT 739
Db      161  HisPheGluAsnIleYsaGlyYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsa 180
QY      740  TTAATAGATGAGAAATTAATCTCTGAACTGAGCTGATTTTATGGGCGAGACTTTGCT 799
Db      181  LeuIleAspGlyGluLeuYsaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 200
QY      800  ATCTTCGAGACTTTGGGCAACCAACCAATCAAGACAGACAGATGATTTCCAGTGG 859
Db      201  IlePheArgThrLeuGlnIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 220
QY      860  CTCATGATCCAAAGTTTCATTAAGTCCACCTCATCTCAGAGAGTGAACATCTGAAAGAT 919
Db      221  LeuAsnaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsa 240
QY      920  GACAAAGTACTTTTCTTCCTCGTGAATATGCAATAGATGAGAACACTCTGGAAAAGCT 979
Db      241  AsplysValIlyrPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 260
QY      980  ACTCAGCTAGAAATAGATGATGATGCAAG 1009
Db      261  ThrHisAlaArgIleGlyGlnIleCyslys 270

RESULT 14
US-09-724-676A-66689
; Sequence 66689, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 66689
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66689

Alignment Scores:
Pred. No.: 4,45e-263 Length: 272
Score: 270.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.17% Indels: 0

```

```

DB:      21  Gaps: 0
US-09-774-490-1 (1-2709) x US-09-724-676A-66689 (1-272)
QY      200  ATGGGCTGTTACTAGAGATGTCTGTCTTTTCTGGGAGATATTAATTAACAGCAAGAGA 259
Db      1  MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValIleuLeuThrAlaArgAla 20
QY      260  AKCTACAAAGGAGAAACAATGTCACAGGCTGAAGTAATTTCTTACAAAGAAATGTTG 319
Db      21  AenlyrglnasngilylsaaasnaValProArgluLeuylseuSerTyrlysglumeLeu 40
QY      320  GAATCCAACAATGATGATCACTTTCATATGCTGGCCCAACAGCTCCAGTTTCATACCTTC 379
Db      41  GluseAsnaValIlethrphenngilyleuAlaasnserserSerTyrhisrhphe 60
QY      380  CTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAATATTTTCATTCATTC 439
Db      61  LeuleuAspGluGluarGserarGleuTyrValGlyAlaYsaPhisrIlepheserPhe 80
QY      440  GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATTTACCAAGAGA 499
Db      81  AspleuValAsnIleuYsaPheGlnlyValIleValIleProValserTyrThrArgArg 100
QY      500  GATGAATGCAAGTGGGCTGGAAGAAGACATCTGAAAGATGTGCTAATTTTCAACAGTA 559
Db      101  AspgIuCylystrpAlagIlysaPrlleuYsaIleuYsaIleuYsaIleuYsaIleuYsa 120
QY      560  CTTAAGGCATATATACAGACTCACTGTACGCTGTGGAACGAGGCTTTTCATCCAAAT 619
Db      121  LeuYsaIaIaYsaGlnThrhisIleuYsaIaIaCysglYThrGlyAlaPhehisProIle 140
QY      620  TGCACTCACTAATTAATGAGCATCATCTGAGAGCAATTTTAACTGGAGAACTCA 679
Db      141  CysThrTyrIleGlnIleGlyhisIshIshIshIshIshIshIshIshIshIshIshIshIsh 160
QY      680  CATTTGAAAACGCGCTGGGAAGAGTCATATGACCTTAAGCTGAGCAGCATCCCTT 739
Db      161  HisPheGluAsnIleYsaGlyYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsa 180
QY      740  TTAATAGATGAGAAATTAATCTCTGAACTGAGCTGATTTTATGGGCGAGACTTTGCT 799
Db      181  LeuIleAspGlyGluLeuYsaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 200
QY      800  ATCTTCGAGACTTTGGGCAACCAACCAATCAAGACAGACAGATGATTTCCAGTGG 859
Db      201  IlePheArgThrLeuGlnIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 220
QY      860  CTCATGATCCAAAGTTTCATTAAGTCCACCTCATCTCAGAGAGTGAACATCTGAAAGAT 919
Db      221  LeuAsnaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsaPrlYsa 240
QY      920  GACAAAGTACTTTTCTTCCTCGTGAATATGCAATAGATGAGAACACTCTGGAAAAGCT 979
Db      241  AsplysValIlyrPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 260
QY      980  ACTCAGCTAGAAATAGATGATGATGCAAG 1009
Db      261  ThrHisAlaArgIleGlyGlnIleCyslys 270

RESULT 15
US-09-864-761-47112
; Sequence 47112, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864.761
; CURRENT FILING DATE: 2001-05-23

```

```

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47112
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
US-09-864-761-47112

Alignment Scores:
Pred. No.: 1,35e-142 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-47112 (1-151)

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTAGTCTACAA 2119
DB 1 IleaTgValaSpaPhisIleIleatgThraSpGInglYleuLeuLeuAArgSerIeuGIn 20
QY 2120 CAGAAGGATTCGGAATTAATCTGCGCCATGCGGCGGAGAACTGGTTCATACAACTCTT 2179
DB 21 GInlySaSpSerGlyAsmlyrIeucYshSaIaValaGluHnSeglyPheIleGInThreU 40
QY 2180 CTTAAGGTAACTCGTAAGTCAATGACACAGACGATTTGGAAGAATCTTCTATAAGT 2239
DB 41 LeuYsValThrIeuGluValIleAspThrGluHnSleuGInGluLeuHnIleAlaYsAp 60
QY 2240 GATGATGAGATGGCTCTAAGCAACAAGAAATGTCATATGATGATACACTTACGAGAG 2299
DB 1 IleaTgValaSpaPhisIleIleatgThraSpGInglYleuLeuLeuAArgSerIeuGIn 20
```

```

DB 61 AapaSpGlyAspGlySerlyeThrlySgluNetSerAenSerMetThrProSerGInlyS 80
QY 2300 GCTGTGATCAGAGATTCATGACGCTCATCAACCAATCTCAACGATGATGAG 2359
DB 81 ValTrpYrArgAspPheMetCInleuLeuHnSProAsnLeuAenThrMetAspGlu 100
QY 2360 TTCTGTGAACAATTGGAGAAAGGAGCAAAACAACGTCGGCAAAAGGCGACATACC 2419
DB 101 PheCySgluGInValTrpIlyAspArgParglySgluAArgGlnaArgProGlyHnSthr 120
QY 2420 CCAGGAAACGTAACAAATGCAAGCATTAACAAGAAATAGAAAGTGAACAGAGG 2479
DB 121 ProGlyAenSerAspIlyeTrpIlyShIleuGInGluAenlySlyGlyAArgAsnArgYr 140
QY 2480 ACCCAGATTTGAGAGGAGCCCGAGAGTGC 2512
DB 141 ThrHISgluPheGluArgAlaProArgSerVal 151

RESULT 16
US-10-203-135-32595
; Sequence 32595, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 5
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US/10/203,135
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 32595
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
US-10-203-135-32595

Alignment Scores:
Pred. No.: 1,35e-142 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-135-32595 (1-151)

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTAGTCTACAA 2119
DB 1 IleaTgValaSpaPhisIleIleatgThraSpGInglYleuLeuLeuAArgSerIeuGIn 20
```

QY 2120 CAGAAGATTACAGCAATTAACCTCTGCGATCCGGGGAACAATGGGTTCATCAAACTCTT 2179
 Db 21 GlnlyAspSerSerIleYasnTrLeuCyshiAlaValGlnHISglYpHeIIegInThrLeu 40
 QY 2180 CTTTAAGTAAACCTTGAAAGTCAATTGACACAGAGCATTTGGAGAACTTCTTCATTAAGAT 2239
 Db 41 LeuLyValThrIeuGlnValIleAspThrGlnHisIeuGlnGlnLeuHISlyAsp 60
 QY 2240 GATGATGAGATGGCTTCTTAAGACCAAGAAATGTCCAAATACATGACACTTAGCCAGAA 2299
 Db 61 AspaSpGlyAspIlySerIlyStrIlySglMetSerAsnSerMetThrProSerGlnlys 80
 QY 2300 GTCTGGTACAGAGACTTCATGACAGCTCATCAACACCCCAATCTCAACAGCATGATGAG 2359
 Db 81 ValTrpTrpArgAspPheMetGlnIleuIleAsnHisProAlaIleuAsnThrMetAspGln 100
 QY 2360 TTCTGTGAACAGATTGGAAAGGAGCCGAAACCAACGTCGGCAAGGCGCAGACATACC 2419
 Db 101 PheCySglGlnGlnValTrpIlyAspArgIlySglIleArgGlnIleArgGlnIlyStrThr 120
 QY 2420 CCAGGGAACAGTAAACAAATGGAGCACTTACACAGAAATAGAAAGGTAGAAACAGAGG 2479
 Db 121 ProGlyAsnSerSerAsnIlyStrIlyShIleGlnGlnIleuSnllySglYlyArgAsnArgArg 140
 QY 2480 ACCCAGCAATTTGAGAGGGCACCAGGAGTGTCTC 2512
 Db 141 ThrHisGlnPheGlnArgAlaProIlyArgSerVal 151
 RESULT 17
 US-10-203-136-33448
 Sequence 33448, Application US/10203136
 GENERAL INFORMATION:
 APPLICANT: Molecular Dynamics, Inc.
 APPLICANT: Penn. Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
 FILE REFERENCE: PB 0004 WO 3
 CURRENT APPLICATION NUMBER: US/10/203,136
 CURRENT FILING DATE: 2002-08-02
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 04 February 2000 (04.02.00)
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 26 May 2000 (26.05.00)
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 03 August 2000 (03.08.00)
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 03 October 2000 (03.10.00)
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 27 September 2000 (27.09.00)
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 21 September 2000 (21.09.00)
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 30 June 2000 (30.06.00)
 NUMBER OF SEQ ID NOS: 38578
 SOFTWARE: Molecular Dynamics Sequence Listing Engine
 SEQ ID NO 33448
 LENGTH: 151
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC006322.2
 FEATURE:
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
 FEATURE:
 OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 2.00e-88
 US-10-203-136-33448
 Alignment Scores:
 1.35e-142 Length: 151
 Pred. No.:

Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-136-33448 (1-151)

OY 2060 ATCAGAGTGGATGATTCATATCATACAGACAGATCAAGGCCCTTGCTAGCTAAGTTCTTCAA 2119
Db 1 ILeArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuAspSerLeuGln 20
OY 2120 CAGAAGATTGAGGCATTTACTCTGCATCGCGGTGAACAATGGGCTTCATAACAACCTT 2179
Db 21 GInLysAspSerGlyAsnTrpLeuCySHisAlaValGlnIleGlyPheIleGlnThrLeu 40
OY 2180 CTTAAGGTAACCCCTGAAAGTCATTTGACACAGAGCATTGGAGAACCTTCTTCATTAAGAT 2239
Db 41 LeuLysValThrLeuGluValIleAspThrGlnHisLeuGluGluLeuLeuHisLysAsp 60
OY 2240 GATGATGAGATGGCTCTTAAGACCAGAAAAGAAATGTCCAATGACATGACACCTTAGCCAGAG 2299
Db 61 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProserGlnLys 80
OY 2300 GTCTGGTACAGAGCTTCATACGACCTCATCAACCCCATACTTCACACAGATGGAAGAG 2359
Db 81 ValTrpLysTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnTrpMetAspGlu 100
OY 2360 TTCTGTGAACAAGTTTGAAAAAGGAGCCGAAAAACAACGTCGGCAAGAGCCAGACATACC 2419
Db 101 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgGlnProGlnHisThr 120
OY 2420 CCAGGGAACAGTAAACAATGGAAGCACTTCAAGAAAATTAAGAAAGGTAGAAAACAGAGG 2479
Db 121 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 140
OY 2480 ACCGACGAATTTGAGAGGGCACCCAGAGTGTC 2512
Db 141 ThrHisGluPheGluArgAlaProArgSerVal 151

RESULT 18
US-09-791-537-72548
; Sequence 72548, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72548
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-72548

Alignment Scores:
Pred. No.: 6,47e-83 Length: 772
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.28% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-72548 (1-772)

OY 1169 GTATATGAGTGTTTACGATCTTCCAGTAACATTTTCAAGGATCAAGCCGCTGATATGAT 1228
Db 324 ValTrpCylValAlaPheThrThrSerSerAsnIlePheLysGlySerIleAlaValCysMetTrp 343

```

QY 1229 ACCATGATGATGTGAGAAAGGTGTTCTTGTTCATATGCCCACAGGATGAGCCCAAC 1288
    |||||
DB 344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
QY 1289 TATCATGGGTGCTTATATCAAGAGATCCCTTATCCAGGCGCAAGAACTGTCCACG 1348
    |||||
DB 364 TyrGlnTyrPValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
QY 1349 AAAACATTTGGTGTGTTGACTCTCAAAAGACCTTCGATGATGTTAATACCTTTGCA 1408
    |||||
DB 384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
QY 1409 AGAAGTCATCCAGCCATGATCAATCCAGTGTTCCT 1444
    |||||
DB 404 ArgSerHisProAlaMetCysTrpAsnProValPhePro 415

RESULT 19
US-09-791-537-101922
; Sequence 101922, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 101922
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-101922

Alignment Scores:
Pred. No.: 6, 47e-83 Length: 772
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.28% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-101922 (1-772)
QY 1169 GTATATGAGTGTTTAAGACTTCCAGTAACTTTTCAGAGATCAGCCGTGTATGAT 1228
    |||||
DB 324 ValTyrGlyValPheThrThrSerSerAsnIlePheLeuGlySerAlaValCysMetCys 343
QY 1229 AGCAGATGATGTGAGAAAGGTGTTCTTGTTCATGATGCCACAGGATGAGCCCAAC 1288
    |||||
DB 344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
QY 1289 TATCATGGGTGCTTATATCAAGAGATCCCTTATCCAGGCGCAAGAACTGTCCACG 1348
    |||||
DB 364 TyrGlnTyrPValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
QY 1349 AAAACATTTGGTGTGTTGACTCTCAAAAGACCTTCGATGATGTTAATACCTTTGCA 1408
    |||||
DB 384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
QY 1409 AGAAGTCATCCAGCCATGATCAATCCAGTGTTCCT 1444
    |||||
DB 404 ArgSerHisProAlaMetCysTrpAsnProValPhePro 415

RESULT 20
US-09-791-537-62757
; Sequence 62757, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek

```

```

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 62757
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-62757

Alignment Scores:
Pred. No.: 8, 67e-73 Length: 666
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.16% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-62757 (1-666)
QY 899 GAGAGTGAACAATCTGAAGATGACAAAGTATATCTTTCTCCGTGAAGAATGCAATAGAT 958
    |||||
DB 128 GluSerAspAsnProGluAspAspValTyrPhePhePheArgGluAsnAlaIleAsp 147
QY 959 GGAGAACACTGTGGAAGAAAGTACTACGCTGAGATAGTGCATATGCAAGATGACTTT 1018
    |||||
DB 148 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 167
QY 1019 GGAGGCGACAGAGCTGCTGAATTAATGACAAATCTCCAAAGCTCGTCTGATTTGC 1078
    |||||
DB 168 GlyGlyHisArgSerLeuValAsnLysTyrThrThrPheLeuLysAlaArgLeuIleCys 187
QY 1079 TCAGTCCAGGCTCAAAATGATGACATCTATTTTGATGAACTGACAGATGATTTCTTA 1138
    |||||
DB 168 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 207
QY 1139 ATGAAC 1144
    |||||
DB 208 MetAsn 209

RESULT 21
US-09-262-167-32
; Sequence 32, Application US/09262167A
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Lofton-Day, Catherine B.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-7
; FILE REFERENCE: 97-59
; CURRENT APPLICATION NUMBER: US/09/262,167A
; EARLIER FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 60/076,611
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-262-167-32

Alignment Scores:
Pred. No.: 9, 35e-70 Length: 772
Score: 79.00 Matches: 79
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.83% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-262-167-32 (1-772)

```



```
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33541
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
PCT-US01-00663-33541
```

Alignment Scores:

Pred. No.:	2,316-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	1	Gaps:	0

US-09-774-490-1 (1-2709) x PCT-US01-00663-33541 (1-57)

```
QY 311 GAATGTTGAATCCAAAGATGATCATCTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 GtmetleugluserasnsvalletmrphasnclyleuAlasnserserxtyr 24
QY 371 CATACCTCTTTGGATGAGAACGAGTAGGCTGTATGTGGAGCAAGATCACAATA 430
DB 25 HsthrPhelenuhspglugluArgserArgleutyValglYAlalyshphtstle 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheserPheaspLeuValasmllelyshphtstle 57
```

RESULT 29

US-09-864-761-44075
Sequence 44075, Application US/09864761

```
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
```

```
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 44075
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
US-09-864-761-44075
```

Alignment Scores:

Pred. No.:	2,316-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	23	Gaps:	0

US-09-774-490-1 (1-2709) x US-09-864-761-44075 (1-57)

```
QY 311 GAATGTTGAATCCAAAGATGATCATCTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 GtmetleugluserasnsvalletmrphasnclyleuAlasnserserxtyr 24
QY 371 CATACCTCTTTGGATGAGAACGAGTAGGCTGTATGTGGAGCAAGATCACAATA 430
DB 25 HsthrPhelenuhspglugluArgserArgleutyValglYAlalyshphtstle 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheserPheaspLeuValasmllelyshphtstle 57
```

RESULT 30
US-10-182-993-32519
Sequence 32519, Application US/10182993


```

GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT APPLICATION NUMBER: US/10/182,993
PRIORITY FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32519
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU E 5.00e-25
US-10-182-993-32519

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 27 Gaps: 0

US-09-774-480-1 (1-2709) x US-10-182-993-32519 (1-57)

OY 311 GAATGTTGGAAATCCCAAGTGATCATTGAATGGCTTGGCGAACAGCTCCAGTTAT 370
Db |||||||
5 GTmettleuglufserahnsavallieturPhaandlyleulaanserSerTyr 24

OY 371 CATACCTTCCTTTGGATGAGGACGAGTAGGCTGTATGTTGGAGCAAAGATCACATA 430
Db |||||||
25 hIstrirPhelLeuaspclungIargSerArgLeutyValGlalalyasphIsile 44

OY 431 TTTTCATGCAGCTGGTTAATATCAAGATTTTCAAAG 469
Db |||||||
45 PheserPhespuenvalasnillelysaspPhedInlys 57

RESULT 31
US-10-203-134-33356
Sequence 33356, Application US/10203134
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 6

```

```

CURRENT APPLICATION NUMBER: US/10/203,134
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33356
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
FEATURE:
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
US-10-203-134-33356

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-134-33356 (1-57)

QY 311 GAAATGTTGGAATCCAAACATGATGATCACTTTCATAGGCTTGCGCAACAGCTCCAGTTAT 370
|||||
DB 5 GlnueclbuGluSerAshmenValIletrPheasnGlyLeuAlaanserSerIyr 24
|||||

QY 371 CATACTTCCTTTTGATGAGGAAACGAGTACGAGCTGTATGTTGAGCAAGATCACATA 430
|||||
DB 25 HisthrPheleuLeuaspGluuIuargserArgleuIyrValGlyAlaIuyspHisIle 44
|||||

QY 431 TTTTCATTGACCTGCTTTAATATCAAGATTTTCAAAAG 469
|||||
DB 45 PheSerPheaspLeuValaInIleIysaspPheGlnIys 57
|||||

RESULT 32
US-10-203-135-32559
Sequence 32559, Application US/10203135
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
FILE REFERENCE: PB 0004 WO 5
CURRENT APPLICATION NUMBER: US/10/203,135
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6

```

```

; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 32559
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
; US-10-203-135-32559

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-135-32559 (1-57)

QY 311 GAAATGTGGAATCCAAATGATGATCACTTCAATGCTTGCCCAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvalillethrphensnglyleuvalaasnserxertyr 24

QY 371 CATACCTTCCTTTGGATGAGGAGGAGTGGCTGTATGTTGGAGCAAGATCACATA 430
DB 25 Hsthrphelenuaspjluglunrgerargleuryvalgllyalysasphstille 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 Pheserpheaspleuvalasnilleysasphneglnlys 57

RESULT 33
US-10-203-136-33412
; Sequence 33412, Application US/10203136
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 3
; CURRENT APPLICATION NUMBER: US/10/203,136
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38578
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
```

```

; SEQ ID NO 33412
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
; US-10-203-136-33412

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-136-33412 (1-57)

QY 311 GAAATGTGGAATCCAAATGATGATCACTTCAATGCTTGCCCAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvalillethrphensnglyleuvalaasnserxertyr 24

QY 371 CATACCTTCCTTTGGATGAGGAGGAGTGGCTGTATGTTGGAGCAAGATCACATA 430
DB 25 Hsthrphelenuaspjluglunrgerargleuryvalgllyalysasphstille 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 Pheserpheaspleuvalasnilleysasphneglnlys 57

RESULT 34
US-10-203-137-33541
; Sequence 33541, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33541
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; FEATURE:
```

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-137-33541

Alignment Scores:

Pred. No.:	2.31e-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	28	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-203-137-33541 (1-57)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTAT 370

DB 5 Gtmetleugluserashnsnvallettrpneannglyleualaansersersertyr 24

QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430

DB 25 HsthrpneuleuaspelugluargerargleutyValGlyAlalyasphistile 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAG 469

DB 45 Pneserpneaspneualasnlelyasppneglmly 57

RESULT 35

US-10-203-139-32331

Sequence 32331, Application US/10203139

GENERAL INFORMATION:

APPLICANT: Molecular Dynamics, Inc.

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: PB 0004 WO 4

CURRENT FILING DATE: 2002-08-02

PRIOR FILING DATE: 04 February 2000 (04.02.00)

PRIOR FILING DATE: 26 May 2000 (26.05.00)

PRIOR FILING DATE: 03 August 2000 (03.08.00)

PRIOR FILING DATE: 03 October 2000 (03.10.00)

PRIOR FILING DATE: 27 September 2000 (27.09.00)

PRIOR FILING DATE: 21 September 2000 (21.09.00)

PRIOR FILING DATE: 30 June 2000 (30.06.00)

NUMBER OF SEQ ID NOS: 37156

SOFTWARE: Molecular Dynamics Sequence Listing Engine

SEQ ID NO 32331

LENGTH: 57

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO ACO04848.1

FEATURE:

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25

US-10-203-139-32331

Alignment Scores:

Pred. No.:	2.31e-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0

DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-139-32331 (1-57)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTAT 370

DB 5 Gtmetleugluserashnsnvallettrpneannglyleualaansersersertyr 24

QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430

DB 25 HsthrpneuleuaspelugluargerargleutyValGlyAlalyasphistile 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAG 469

DB 45 Pneserpneaspneualasnlelyasppneglmly 57

RESULT 36

US-10-029-386-29372

Sequence 29372, Application US/10029386

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: AEWICA-X-2

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 29372

LENGTH: 60

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR7.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.68

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.69

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25

US-10-029-386-29372

Alignment Scores:

Pred. No.:	2.3e-43	Length:	60
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	26	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-029-386-29372 (1-60)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTAT 370

DB 4 Gtmetleugluserashnsnvallettrpneannglyleualaansersersertyr 23

QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430

DB 24 HsthrpneuleuaspelugluargerargleutyValGlyAlalyasphistile 43

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAG 469

DB 44 Pneserpneaspneualasnlelyasppneglmly 56

RESULT 37

US-09-500-746-13

Sequence 13, Application US/09500746

GENERAL INFORMATION:

APPLICANT: Winchester, Robert J.

APPLICANT: Guiko, Percio

APPLICANT: Seki, Tetsumori

```

/ TITLE OF INVENTION: USFS OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
/ FILE REFERENCE: 0575/57005-B
/ CURRENT APPLICATION NUMBER: US/09/500,746
/ CURRENT FILING DATE: 2000-02-09
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 13
/ LENGTH: 101
/ TYPE: PRT
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (54) (54)
/ OTHER INFORMATION: X=to any amino acid
US-09-500-746-13

Alignment Scores:
Pred. No.: 2,23e-43 Length: 101
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-500-746-13 (1-101)
OY 1211 TCAGCGTGTGTATGATAGCATGAGTGTGAGAAAGGTGTTCTTGTCATATGCC 1270
DB 1 SerAlaValCymetyrisermetserAspValArgValPheIenGlyProTyrAla 20
OY 1271 CACAGGATGAGCCCACTATCATATGGTGCCCTTATCAAGAAAGATCCCCATTCACGG 1330
DB 21 HisArgAspGlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArg 40
OY 1331 CCAGGAAGCTGTCCAGCAAAACATTGGTGTTTGAC 1369
DB 41 ProGlyThrCysProSerLysThrPheGlyGlyPheAsp 53

RESULT 38
US-60-160-203-3958
/ Sequence 1958, Application US/60160203
/ GENERAL INFORMATION:
/ APPLICANT: BONAZZI, VIVIAN
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
/ TITLE OF INVENTION: USFS THEREOF
/ FILE REFERENCE: CLO00116
/ CURRENT APPLICATION NUMBER: US/60/160,203
/ CURRENT FILING DATE: 1999-10-19
/ NUMBER OF SEQ ID NOS: 6374
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3958
/ LENGTH: 50
/ TYPE: PRT
/ ORGANISM: HUMAN
US-60-160-203-3958

Alignment Scores:
Pred. No.: 2,69e-38 Length: 50
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-160-203-3958 (1-50)
OY 866 GATCCAAAGTTCATTATAGTCCACCACTCATCTGAGAGTGAACAATCTGAGATGACAA 925
DB 3 AspProLysPheIleSerAlaHisLeuIleSerGlnSerAspAsnProGluAspHis 22
OY 926 GTATACCTTTTCTCCGGAATGCATATGATGAGAAACACTCTGGAAAACCTACTCAC 985

```

```

Db      23 Valtyrhehehehargluamalaileapglglnhisetcdlyysalathrhls 42
OY      986 GCTGGAATAGCTGAGATATGCAAG 1009
Db      43 Alaargileglglnilecyslys 50

RESULT 39
US-60-160-203-4025
; Sequence 4025, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONNAZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4025
; LENGTH: 50
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4025

Alignment Scores:
Pred. No.: 2,69e-38 Length: 50
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-160-203-4025 (1-50)
OY      866 GATCCAAAGTTCATTGATGCGCCACCTCATCTGAGAGAGACATCTGGAAGATGACAA 925
Db      3 AspprolysheheleaserlahisleulleasergluserAspamProdluspaplys 22
OY      926 GTTACTCTTTCTTCGGTGAATAATGCATAGATGAGAACTCTGGAAGAAAGCTACTCAC 985
Db      23 Valtyrhehehehargluamalaileapglglnhisetcdlyysalathrhls 42
OY      986 GCTGGAATAGCTGAGATATGCAAG 1009
Db      43 Alaargileglglnilecyslys 50

RESULT 40
US-09-864-761-44552
; Sequence 4452, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hankel, David R.
; APPLICANT: Hankel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 44552
;; LENGTH: 41
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
;; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)
```

Alignment Scores:

Pred. No.:	3,32e-31	Length:	41
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.58%	Indels:	0
DB:	23	Gaps:	0

```
US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)
QY 1937 CCTGAAGAGAGATCATGTATGTTAGAGATGATGACATTTTGGATGAGTCCG 1996
DB 1 ProgluGluArgIleIleIleTyrlGlyValGluAnsSerSerrPheLeuGluCysSerPro 20
QY 1997 AAGTCGACAGAGCGCTGTCTATTGCAATTCAGAGCGGCAATGAAGCGAAGA 2056
DB 21 LysSerGlnArgAlaLeuValTyrrTgInPheGlnArgArgAnsGluGluArgLysGlu 40
QY 2057 GAG 2059
DB 41 Gln 41
RESULT 41
US-10-182-993-33490
;; Sequence 33490, Application US/10182993
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN BRAIN
;; FILE REFERENCE: PB 0004 WO 2
```

```
;; CURRENT APPLICATION NUMBER: US/10/182,993
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37811
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 33490
;; LENGTH: 41
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
;; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-10-182-993-33490
```

Alignment Scores:

Pred. No.:	3,32e-31	Length:	41
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.58%	Indels:	0
DB:	27	Gaps:	0

```
US-09-774-490-1 (1-2709) x US-10-182-993-33490 (1-41)
QY 1937 CCTGAAGAGAGATCATGTATGTTAGAGATGATGACATTTTGGATGAGTCCG 1996
DB 1 ProgluGluArgIleIleIleTyrlGlyValGluAnsSerSerrPheLeuGluCysSerPro 20
QY 1997 AAGTCGACAGAGCGCTGTCTATTGCAATTCAGAGCGGCAATGAAGCGAAGA 2056
DB 21 LysSerGlnArgAlaLeuValTyrrTgInPheGlnArgArgAnsGluGluArgLysGlu 40
QY 2057 GAG 2059
DB 41 Gln 41
RESULT 42
US-10-203-135-33728
;; Sequence 33728, Application US/10203135
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
;; FILE REFERENCE: PB 0004 WO 5
;; CURRENT APPLICATION NUMBER: US/10/203,135
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
```

;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37012
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 33728
;; LENGTH: 41
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
;; FEATURE:
;; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
;; FEATURE:
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-10-203-135-33728

Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-10-203-135-33728 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGAGAGATAGTGCACATTTTGGATCGACTCG 1996
|||
DB 1 ProgluinarigileiletyrGlyValGluAnserserThrPheleuglucysserPro 20

QY 1997 AAGTCGAGAGAGCGGTGCTATTGGCAATTCAGAGCGGGAATGAAGCGGAAGA 2056
|||
DB 21 LysSerGlnArgAlaLeuValItyrTrpGlnPheGlnArgAlaGlnGluArgLysGlu 40

QY 2057 GAG 2059
|||
DB 41 Glu 41

RESULT 43
US-10-203-136-34569
;; Sequence 34569, Application US/10203136
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn. Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
;; FILE REFERENCE: PB 0004 WO 3
;; CURRENT APPLICATION NUMBER: US/10/203,136
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)

;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38578
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 34569
;; LENGTH: 41
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
;; FEATURE:
;; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
;; FEATURE:
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-10-203-136-34569

Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-10-203-136-34569 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGAGAGATAGTGCACATTTTGGATCGACTCG 1996
|||
DB 1 ProgluinarigileiletyrGlyValGluAnserserThrPheleuglucysserPro 20

QY 1997 AAGTCGAGAGAGCGGTGCTATTGGCAATTCAGAGCGGGAATGAAGCGGAAGA 2056
|||
DB 21 LysSerGlnArgAlaLeuValItyrTrpGlnPheGlnArgAlaGlnGluArgLysGlu 40

QY 2057 GAG 2059
|||
DB 41 Glu 41

RESULT 44
US-60-177-571-3274
;; Sequence 3274, Application US/60177571
;; GENERAL INFORMATION:
;; APPLICANT: Bonazzi, Vivien
;; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
;; FILE REFERENCE: CL000201
;; CURRENT APPLICATION NUMBER: US/60/177,571
;; CURRENT FILING DATE: 2000-01-27
;; NUMBER OF SEQ ID NOS: 5082
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3274
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: HUMAN
US-60-177-571-3274

Alignment Scores:
Pred. No.: 3,55e-28 Length: 53
Score: 38.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.25% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-60-177-571-3274 (1-53)

QY 1916 GATATGACCATGCGGCAAGCCCTGAAGAGAGATCATCTATGCTGAGAGATAGTACC 1975
|||
DB 5 AspaenH18H18GlyHisSerProGluIuArgIleItyrGlyValGluAnserser 24

QY 1976 ACATTTTGCATGCTCCGAAAGTCCAGAGAGCCCTGCTATTGGCAATTC 2029
DB 25 ThrPheLeuGlnGlySerProGlySerGlnArgAlaLeuValTyrTrpGlnPhe 42

RESULT 45
US-09-500-746-9
Sequence 9, Application US/09500746
GENERAL INFORMATION:
APPLICANT: Winchester, Robert J.
APPLICANT: Gulko, Percio
APPLICANT: Seki, Tetsunori
TITLE OF INVENTION: USES OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
FILE REFERENCE: 0575/57005-B
CURRENT APPLICATION NUMBER: US/09/500,746
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 507
TYPE: PRT
ORGANISM: mouse
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (337)..(337)
OTHER INFORMATION: x= to any amino acid
NAME/KEY: MISC FEATURE
LOCATION: (376)..(376)
OTHER INFORMATION: x= to any amino acid
US-09-500-746-9

Alignment Scores:
Pred. No.: 3,34e-25 Length: 507
Score: 35.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-500-746-9 (1-507)

QY 1241 GTGAGAGGGGTTCCTTGGTCATATGCGCCAGAGGATGACCCCACTATCAATGGGTG 1300
DB 174 ValArgGlyValPheLeuGlnGlyProTyrAlaHisArgGlyProAsnTyrGlnTrpVal 193

QY 1301 CCTTCAAGAGAGTCCCTATCCAGCGCCAGGAACCTTGTCCC 1345
DB 194 ProTyrGlnGlyArgValProTyrProArgProGlyThrCysPro 208

RESULT 46
PCT-US01-00663-33648
Sequence 33648, Application PC/TUS0100663
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: US 60/180,312
PRIORITY FILING DATE: 04 February 2000 (04.02.00)
PRIORITY FILING DATE: US 60/207,456
PRIORITY FILING DATE: 26 May 2000 (26.05.00)
PRIORITY FILING DATE: US 09/632,366
PRIORITY FILING DATE: 03 August 2000 (03.08.00)
PRIORITY FILING DATE: GB 24263.6
PRIORITY FILING DATE: 03 October 2000 (03.10.00)
PRIORITY FILING DATE: US 60/236,359
PRIORITY FILING DATE: 27 September 2000 (27.09.00)

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33648
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION 4.00e-15
PCT-US01-00663-33648

Alignment Scores:
Pred. No.: 4,22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US01-00663-33648 (1-33)

QY 533 AAGAGATGCTAAATTCATCAGGTACTTAAGCATTAATCAGACTCTTGTACGCC 592
DB 1 LysGlnCysAlaAsnPheIleValLeuValAlaTyrAsnGlnThrHisLeuTyrAla 20

QY 593 TGTGAGCGGGGCTTTTCATCCATTTGACACTTACATT 631
DB 21 CysGlnThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 47
US-09-864-761-39708
Sequence 39708, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIORITY FILING DATE: 2001-05-23
PRIORITY FILING DATE: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY FILING DATE: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY FILING DATE: 2000-10-04
PRIORITY FILING DATE: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY FILING DATE: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39708
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-09-761-39708

Alignment Scores:
Pred. No.: 4.22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-39708 (1-33)

QY 533 AAGAATGCTAATTCATCAAGTACTTAAGCATATATACAGCTCACTGTAGGCC 592
|||||
DB 1 LysGluCyAlaAsnPhelIeYsValLeuYsAlaIryAsnGlnThrHisLeuYrAla 20

QY 593 TGTGAAACGGGGCTTTTCATTCACATTTGCACCTACATT 631
|||||
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 48
US-10-182-993-32618

Sequence 32618, Application US/10182993
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn. Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT APPLICATION NUMBER: US/10/182,993
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32618
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-10-182-993-32618

Alignment Scores:
Pred. No.: 4.22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 27 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-182-993-32618 (1-33)

QY 533 AAGAATGCTAATTCATCAAGTACTTAAGCATATATACAGCTCACTGTAGGCC 592
|||||
DB 1 LysGluCyAlaAsnPhelIeYsValLeuYsAlaIryAsnGlnThrHisLeuYrAla 20

QY 593 TGTGAAACGGGGCTTTTCATTCACATTTGCACCTACATT 631
|||||
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 49
US-10-182-995-26180

Sequence 26180, Application US/10182995
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn. Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 1
CURRENT APPLICATION NUMBER: US/10/182,995
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
NUMBER OF SEQ ID NOS: 29119
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 26180
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15

US-10-182-995-26180

Alignment Scores:

Pred. No.:	4,22e-23	Length:	33
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	27	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-182-995-26180 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

Db 1 LysglucyalaasphenlelevalleuysalatyrAsnglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

Db 21 CysglythnrglyAlaPhehisProleleCystnryrile 33

RESULT 50

US-10-203-134-33472

; Sequence 33472, Application US/10203134

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: PB 0004 WO 6

; CURRENT FILING DATE: 2002-08-02

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 38628

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 33472

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC004848.1

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION 4.00e-15

; US-10-203-134-33472

Alignment Scores:	
Pred. No.:	4,22e-23
Score:	33.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.69%
DB:	28

	Length:	33
	Matches:	33
	Conservative:	0
	Mismatches:	0
	Indels:	0
	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-203-134-33472 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

Db 1 LysglucyalaasphenlelevalleuysalatyrAsnglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

Db 21 CysglythnrglyAlaPhehisProleleCystnryrile 33

RESULT 51

US-10-203-135-32678

; Sequence 32678, Application US/10203135

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: PB 0004 WO 5

; CURRENT FILING DATE: 2002-08-02

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 37012

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 32678

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC004848.1

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION 4.00e-15

; US-10-203-135-32678

Alignment Scores:	
Pred. No.:	4,22e-23
Score:	33.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.69%
DB:	28

	Length:	33
	Matches:	33
	Conservative:	0
	Mismatches:	0
	Indels:	0
	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-203-135-32678 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

Db 1 LysglucyalaasphenlelevalleuysalatyrAsnglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

Db 21 CysglythnrglyAlaPhehisProleleCystnryrile 33

RESULT 52

US-10-203-136-33531

; Sequence 33531, Application US/10203136

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

```

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 3
; CURRENT APPLICATION NUMBER: US/10/203,136
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38578
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33531
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 4.00e-15
; US-10-203-136-33531

Alignment Scores:
Pred. No.: 4,22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-136-33531 (1-33)

QY 533 AAGAGATGCTTAATTCATCAAGTACTTAAGCATATATACACTGTTAGCC 592
DB 1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
QY 593 TGTGAACGGGGCTTTTCATCCATTTGCACCTACATT 631
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 53
US-10-203-137-33648
; Sequence 33648, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
```

```

; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33648
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 4.00e-15
; US-10-203-137-33648

Alignment Scores:
Pred. No.: 4,22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-137-33648 (1-33)

QY 533 AAGAGATGCTTAATTCATCAAGTACTTAAGCATATATACACTGTTAGCC 592
DB 1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
QY 593 TGTGAACGGGGCTTTTCATCCATTTGCACCTACATT 631
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 54
US-10-203-139-32427
; Sequence 32427, Application US/10203139
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 4
; CURRENT APPLICATION NUMBER: US/10/203,139
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37156
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 32427
; LENGTH: 33
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; FEATRE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-10-203-139-32427

Alignment Scores:
Pred. No.: 4.22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-139-32427 (1-33)

OY 533 AAAGATGTCATTAATTCATCAAGGTAAGGATTAATCACTTGTACGCC 592
DB 1 LysgluCySalasphelleyValuylSalatyrAsnGlnThiLeuTyrla 20
OY 593 TGTGGAACGGGGCTTTTCATCAATTTGCACCTACATT 631
DB 21 CysglyThrGlyAlaPheHisproIleCysThrTyrlle 33

RESULT 55
US-60-177-571-3503
; Sequence 3503, Application US/60177571
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
; FILE REFERENCE: CLO00201
; CURRENT APPLICATION NUMBER: US/60/177,571
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 5082
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3503
; LENGTH: 50
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-571-3503

Alignment Scores:
Pred. No.: 4.11e-23 Length: 50
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-571-3503 (1-50)

OY 320 GAATCAACATGATCATCTTCAATGCTTGCCACACAGTCCAGTTATCATACCTTC 379
DB 1 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrlHisThrPhe 20
OY 380 CTTTGGATGAGGAGGAGGAGGCTGATGTTGGAGCA 418
DB 21 LeuLeuAspGluGluArgSerArgLeuTyrlValGlyAla 33

RESULT 56
US-60-182-467-1352
; Sequence 1352, Application US/60182467
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
```

```

; FILE REFERENCE: CLO00248
; CURRENT APPLICATION NUMBER: US/60/182,467
; CURRENT FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 2194
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1352
; LENGTH: 60
; TYPE: PRT
; ORGANISM: HUMAN
US-60-182-467-1352

Alignment Scores:
Pred. No.: 4.06e-23 Length: 60
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-182-467-1352 (1-60)

OY 1217 GTGTGATGATATAGATGATGATGAGGAGGTCTTCTTGTCCATATGCCACAG 1276
DB 28 ValCysMetTyrlSerMetSerAspValArgValPheLeuGlyProTyrlAlaHisArg 47
OY 1277 GATGACCCCACTATCAATGGGTGCTTATCAAGGAAGA 1315
DB 48 AspGlyProAsnTyrlGlnTrpValProTyrlGlnGlyArg 60

RESULT 57
US-60-182-467-1353
; Sequence 1353, Application US/60182467
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO00248
; CURRENT APPLICATION NUMBER: US/60/182,467
; CURRENT FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 2194
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1353
; LENGTH: 60
; TYPE: PRT
; ORGANISM: HUMAN
US-60-182-467-1353

Alignment Scores:
Pred. No.: 4.06e-23 Length: 60
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-182-467-1353 (1-60)

OY 1217 GTGTGATGATATAGATGATGATGAGGAGGTCTTCTTGTCCATATGCCACAG 1276
DB 28 ValCysMetTyrlSerMetSerAspValArgValPheLeuGlyProTyrlAlaHisArg 47
OY 1277 GATGACCCCACTATCAATGGGTGCTTATCAAGGAAGA 1315
DB 48 AspGlyProAsnTyrlGlnTrpValProTyrlGlnGlyArg 60

RESULT 58
US-60-163-123-1597
; Sequence 1597, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
```

```

: TITLE OF INVENTION: US$ THEREOF
: FILE REFERENCE: CL000137
: CURRENT APPLICATION NUMBER: US/60/163,123
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 1986
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1597
: LENGTH: 35
: TYPE: PRT
: ORGANISM: Human
US-60-163-123-1597

```

Alignment Scores:	
Pred. No.:	4,456-21
Score:	31.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3,46*
DB:	31
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
	Matches: 31
	Length: 35

US-09-774-490-1 (1-2709) x US-60-163-123-1597 (1-35)

QY 1555 GAGCTTGGGACCGTCTCTTAAGTGTTCATTCCTAAGAGACTTGGATGATTTAGAA 1618
db 1 AspValGlyThrValLeuYsValSerIlePolycystinrrTPYraePleuGu 20

```

Oy      1619 GAGGTTCTGCTGGAGAAATGACAGTTTTTCGG 1651
          |||||
Db      21  GluValLeuGluGluMetThrValPheArg 31

```

RESULT 59
US-09-791-537-105918
; Sequence 105918, Application US/09791537

```

? GENERAL INFORMATION:
? APPLICANT: Biomimix, Inc.
? APPLICANT: Debe, Derek
? APPLICANT: Danzer, Joseph
? TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
? FILE OF INVENTION: METHODS OF USE THEREOF
? FILE REFERENCE: 261/210
? CURRENT APPLICATION NUMBER: US/09/791,537
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 153055
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 105918

```

```

; TYPE: PRT
; ORGANISM: Danio rerio
US-09-791-537-105918

```

Alignment Scores:	
Pred. No.:	4.74e-14
Score:	24.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.68%
DB:	22
Length:	778
Matches:	24
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) x US-09-791-537-105918 (1-778)

Qy 123 TGTTCGCTATTTCCACTGCAGAGAGACGCAAGACGACAAGATTAAAGAAATGGA 1882

Db 543 CysseAaGTYrPhePcThraAlaLysaGArGThrAaGArGAlnaSpLLeaGasnGly 562

```

QY      1883 GACCCACTGACT 1894
      .      |||||
Db      563 AspProLeuThr 566

```

RESULT 60
US-09-513-999C-6632
; Sequence 6632, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.

```

? APPLICANT: Duclert, A.
? APPLICANT: Giordano, J.Y.
? TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
? FILE REFERENCE: 59.US2.REG
? CURRENT APPLICATION NUMBER: US/09/513,999C
? CURRENT FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/122,487
? PRIOR FILING DATE: 1999-02-26
? NUMBER OF SEQ ID NOS: 36681
? SOFTWARE: Patent.pm
? SEQ ID NO 6632

```

```

? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: UNSURE
? LOCATION: 28
? OTHER INFORMATION: Xaa=Glu or Lys
? FEATURE:
? NAME/KEY: UNSURE
? LOCATION: 29

```

```

; OTHER INFORMATION: Xaa=Ala or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 30
;

```

```

;-----;
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 32
;-----;

```

US-09-513-999C-6632

Alignment Scores:

Pred. No.:	5.51e-12	Length:	67
Score:	22.00	Matches:	35
Percent Similarity:	97.22%	Conservative:	0
Best Local Similarity:	97.22%	Mismatches:	0
Query Match:	2.46%	Indels:	1
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x US-09-513-999C-6632 (1-67)

QY 176 AAAGGACCTACAGCGTCTCCAGCATGGGCTGGTTAACTAGAGATTGTCGTCTTTTCTGG 235
DB 33 LysGlyThrTyrSerValCysSerMetGlyTrpLeuThrArgIleValCysLeuPheTrp 52

QY 236 GGAGTATTACTTACAGCAAGCAAACCTATTCGAA TGGGAAGAACA 281
| | | | | | | | | | | | | | | | | | | | | |
Db 53 -GlutyrtyrleuglnglmglnThrIleargmetGIyArgThr 67

RESULT 61
PCT-US00-

; Sequence 16, Application PC/TUS0041943

GENERAL INFORMATION:

APPLICANT: The General Hospital Corporation

FILE REFERENCE: 00786-441W01

FILE REFERENCE: 00786-441WO1
CURRENT APPLICATION NUMBER: PCT/US00/41943

CURRENT FILING DATE: 2000-11-07

PRIOR APPLICATION NUMBER: US 09/694,085

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: US 60/164,056

PRIOR FILING DATE: 1999-11-08

```

; NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```

```

: SEQ ID NO 16
:
: LENGTH: 17

```

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: computer-generated am

PCT-US00-41943-16

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-16 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAT 1813
|||||

Db 1 LysAlaCyAlaGluCyCysLeuAlaArgAspProTyrCyAlaIatrpasp 17

RESULT 62
PCT-US00-41943-43
; Sequence 43, Application PC/TUS0041943
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441M01
; CURRENT APPLICATION NUMBER: PCT/US00/41943
; CURRENT FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/694,085
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: computer-generated amino acid
PCT-US00-41943-43

Alignment Scores:

Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-43 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAT 1813
|||||

Db 1 LysAlaCyAlaGluCyCysLeuAlaArgAspProTyrCyAlaIatrpasp 17

RESULT 63

US-09-694-085-16
; Sequence 16, Application US/09694085
; GENERAL INFORMATION:
; APPLICANT: Behar, Oded
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441001
; CURRENT APPLICATION NUMBER: US/09/694,085
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: computer-generated amino acid
US-09-694-085-16

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-694-085-16 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAT 1813
|||||

Db 1 LysAlaCyAlaGluCyCysLeuAlaArgAspProTyrCyAlaIatrpasp 17

RESULT 64
US-09-694-085-43
; Sequence 43, Application US/09694085
; GENERAL INFORMATION:
; APPLICANT: Behar, Oded
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441001
; CURRENT APPLICATION NUMBER: US/09/694,085
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: computer-generated amino acid
US-09-694-085-43

Alignment Scores:

Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-694-085-43 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAT 1813
|||||

Db 1 LysAlaCyAlaGluCyCysLeuAlaArgAspProTyrCyAlaIatrpasp 17

RESULT 65

US-60-177-571-3765
; Sequence 3765, Application US/60177571
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NOCTERIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
; FILE REFERENCE: CL000201
; CURRENT APPLICATION NUMBER: US/60/177,571
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 5082
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3765
; LENGTH: 73
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-571-3765

Alignment Scores:
Pred. No.: 6.31e-07 Length: 73
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-571-3765 (1-73)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 28 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 44

RESULT 66

US-60-177-646-3105
; Sequence 3105, Application US/60177646
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CL000210
; CURRENT APPLICATION NUMBER: US/60/177,646
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 4226
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3105
; LENGTH: 73
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-646-3105

Alignment Scores:

Pred. No.: 6.31e-07 Length: 73
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-646-3105 (1-73)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 28 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 44

RESULT 67

US-60-177-571-3766
; Sequence 3766, Application US/60177571
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
; FILE REFERENCE: CL000201
; CURRENT APPLICATION NUMBER: US/60/177,571
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 5082
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3766
; LENGTH: 79
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-571-3766

Alignment Scores:

Pred. No.: 6.28e-07 Length: 79
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-571-3766 (1-79)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 26 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 42

RESULT 68

PCT-US00-33116-2
; Sequence 2, Application PC/TUS0033116
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
; FILE REFERENCE: 99-90PC
; CURRENT APPLICATION NUMBER: PCT/US00/33116
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 90/455,560
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US00-33116-2

Alignment Scores:

Pred. No.: 5.42e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-33116-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 525 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 69

US-09-455-560-2
; Sequence 2, Application US/09455560
; GENERAL INFORMATION:
; APPLICANT: Holloway, James E.
; APPLICANT: Foley, Kevin P.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
; FILE REFERENCE: 99-90X
; CURRENT APPLICATION NUMBER: US/09/455,560
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-455-560-2

Alignment Scores:

Pred. No.: 5.42e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 18 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-455-560-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 525 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 70

US-09-731-179-2
; Sequence 2, Application US/09731179
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L..


```
/ APPLICANT: Foley, Kevin P.
/ TITLE OF INVENTION: HUMAN SEMAPHORIN ZSNF-16
/ FILE REFERENCE: 99-90
/ CURRENT APPLICATION NUMBER: US/09/731,179
/ PRIOR FILING DATE: 2000-12-06
/ PRIOR APPLICATION NUMBER: US 60/169,238
/ PRIOR FILING DATE: 1999-12-06
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 779
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-731-179-2

Alignment Scores:
Pred. No.: 5,42e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 21 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-731-179-2 (1-779)
QY 1766 GCGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspIly 544

RESULT 71
PCT-US00-04340-18
/ Sequence 18, Application PC/TUS0004340
/ GENERAL INFORMATION:
/ APPLICANT: Valenzuela, Dario
/ APPLICANT: Yuan, Olive
/ APPLICANT: Hoffman, Heidi
/ APPLICANT: Hall, Jeff
/ APPLICANT: Radtje, Peter
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: GI 6918X
/ CURRENT APPLICATION NUMBER: PCT/US00/04340
/ CURRENT FILING DATE: 2000-02-18
/ NUMBER OF SEQ ID NOS: 101
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 18
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
PCT-US00-04340-18

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-04340-18 (1-782)
QY 1766 GCGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspIly 544

RESULT 72
US-09-507-209-18
/ Sequence 18, Application US/09507209
/ GENERAL INFORMATION:
/ APPLICANT: Valenzuela, Dario
/ APPLICANT: Yuan, Olive
/ APPLICANT: Hoffman, Heidi
/ APPLICANT: Hall, Jeff
/ APPLICANT: Radtje, Peter
```

```
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: GI 6918X
/ CURRENT APPLICATION NUMBER: US/09/507,209
/ PRIOR FILING DATE: 2000-02-18
/ NUMBER OF SEQ ID NOS: 101
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 18
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-507-209-18

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-507-209-18 (1-782)
QY 1766 GCGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspIly 544

RESULT 73
US-09-813-290-4
/ Sequence 4, Application US/09813290
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Milganowski, Nathaniel L.
/ APPLICANT: Turner, C. Alexander Jr.
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Wang, Xiaoming
/ APPLICANT: Donoho, Gregory
/ APPLICANT: Scoville, John
/ TITLE OF INVENTION: Novel Human Secreted Proteins and Polynucleotides Encoding the S
/ FILE REFERENCE: LEX-0151-USA
/ CURRENT APPLICATION NUMBER: US/09/813,290
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: US 60/190,638
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: US 60/191,188
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: US 60/193,639
/ PRIOR FILING DATE: 2000-03-31
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-813-290-4

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-4 (1-782)
QY 1766 GCGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspIly 544

RESULT 74
PCT-US00-04340-92
/ Sequence 92, Application PC/TUS0004340
/ GENERAL INFORMATION:
```

APPLICANT: Valenzuela, Dario
APPLICANT: Yuan, Olive
APPLICANT: Hoffman, Heidi
APPLICANT: Hall, Jeff
APPLICANT: Rajciko, Peter
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6918X
CURRENT APPLICATION NUMBER: PCT/US00/04340
CURRENT FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 92
LENGTH: 785
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-04340-92

Alignment Scores:
Pred. No.: 5,41e-07 Length: 785
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-04340-92 (1-785)

QY 1766 GCCTGCTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 531 AAcYsAlaGluCysCysLeuAlaArgAspProtyrCysAlaTrpAspGly 547

RESULT 75

US-09-507-209-92
Sequence 92, Application US/09507209
GENERAL INFORMATION:
APPLICANT: Valenzuela, Dario
APPLICANT: Yuan, Olive
APPLICANT: Hoffman, Heidi
APPLICANT: Hall, Jeff
APPLICANT: Rajciko, Peter
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6918X
CURRENT APPLICATION NUMBER: US/09/507,209
CURRENT FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 92
LENGTH: 785
TYPE: PRT
ORGANISM: Homo sapiens
US-09-507-209-92

Alignment Scores:
Pred. No.: 5,41e-07 Length: 785
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-507-209-92 (1-785)

QY 1766 GCCTGCTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 531 AAcYsAlaGluCysCysLeuAlaArgAspProtyrCysAlaTrpAspGly 547

Search completed: August 3, 2003, 10:47:06
Job time : 459 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.r2p model

Run on: August 3, 2003, 10:34:54 ; Search time 10 Seconds
(without alignments)
6791.454 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aatcttcatctatcatgcatg.....agcctttcttccataacc 2709

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 67526 segs, 12534983 residues

Word size: 1

Total number of hits satisfying chosen parameters: 128650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+.r2p.model -DEV=xlh
-O=/cgn2_1/USFTO_spool/US09774490/runat_03082003_102929_4545/app_query.fast_1.2887
-DB=pendig.patents.AA.New -OPMT=faetan -SUFIX=oligo.r2pn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USRR=US09774490 @CCN 1.1.9 @runat_03082003_102929_4545
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA.New.*
2: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	15	1.7	514	6	US-10-357-820-28
2	15	1.7	636	6	US-10-408-765A-2003
3	13	1.5	775	1	PCT-US03-09929-92
4	13	1.5	775	1	PCT-US03-09929-94
5	13	1.5	775	1	PCT-US03-09929-96
6	13	1.5	775	1	PCT-US03-09929-98
7	11	1.2	431	6	US-10-369-072-97
8	10	1.1	495	1	PCT-US03-09929-4
9	10	1.1	634	1	PCT-US03-09929-8
10	10	1.1	638	1	PCT-US03-09929-8
11	10	1.1	638	1	PCT-US03-09929-10

12	10	1.1	833	1	PCT-US03-09929-6	Sequence 6, Appl
13	10	1.1	833	1	PCT-US03-09929-12	Sequence 12, Appl
14	10	1.1	833	6	US-10-369-072-14	Sequence 14, Appl
15	10	1.1	833	6	US-10-369-072-16	Sequence 16, Appl
16	10	1.1	833	6	US-10-369-072-18	Sequence 18, Appl
17	10	1.1	834	6	US-10-369-072-47	Sequence 47, Appl
18	10	1.1	963	6	US-10-369-072-46	Sequence 46, Appl
19	10	1.1	350	6	US-10-376-397A-42	Sequence 42, Appl
20	9	1.0	510	6	US-10-369-072-48	Sequence 48, Appl
21	9	1.0	782	6	US-10-369-072-39	Sequence 39, Appl
22	9	1.0	893	6	US-10-369-072-38	Sequence 38, Appl
23	9	1.0	1244	6	US-10-326-956-1715	Sequence 1715, Ap
24	9	1.0	3647	6	US-10-408-765A-2436	Sequence 2436, Ap
25	8	0.9	312	7	US-60-487-610-1445	Sequence 1645, Ap
26	8	0.9	754	6	US-10-292-798-748	Sequence 748, Ap
27	8	0.9	754	6	US-10-369-072-10	Sequence 10, Appl
28	8	0.9	756	6	US-10-369-072-8	Sequence 8, Appl
29	8	0.9	761	6	US-10-369-072-36	Sequence 36, Appl
30	8	0.9	761	6	US-10-293-244-3947	Sequence 3947, Ap
31	8	0.9	762	6	US-10-369-072-35	Sequence 35, Appl
32	8	0.9	838	6	US-10-369-072-49	Sequence 49, Appl
33	8	0.9	1074	6	US-10-273-573-9338	Sequence 9338, Ap
34	8	0.9	1074	6	US-10-369-072-41	Sequence 41, Appl
35	8	0.9	1074	6	US-10-369-072-44	Sequence 44, Appl
36	8	0.9	1077	6	US-10-369-072-43	Sequence 43, Appl
37	8	0.9	5032	6	US-10-408-765A-26	Sequence 26, Appl
38	7	0.8	10	6	US-10-462-850-2385	Sequence 2385, Ap
39	7	0.8	10	6	US-10-462-850-2387	Sequence 2387, Ap
40	7	0.8	10	6	US-10-462-850-2399	Sequence 2399, Ap
41	7	0.8	10	6	US-10-462-850-2401	Sequence 2401, Ap
42	7	0.8	10	6	US-10-462-850-2485	Sequence 2485, Ap
43	7	0.8	10	6	US-10-462-850-2527	Sequence 2527, Ap
44	7	0.8	10	6	US-10-462-850-2533	Sequence 2533, Ap
45	7	0.8	10	6	US-10-462-850-2543	Sequence 2543, Ap
46	7	0.8	10	6	US-10-462-850-2557	Sequence 2557, Ap
47	7	0.8	12	6	PCT-US02-39397-69	Sequence 156, App
48	7	0.8	22	1	PCT-US02-39397-69	Sequence 90, Appl
49	7	0.8	47	6	US-10-603-114-7828	Sequence 7994, Ap
50	7	0.8	72	6	US-10-374-780A-2152	Sequence 2152, Ap
51	7	0.8	121	6	US-10-273-573-7304	Sequence 1204, Ap
52	7	0.8	126	6	US-10-374-780A-1248	Sequence 1248, Ap
53	7	0.8	151	6	US-10-374-780A-1342	Sequence 1342, Ap
54	7	0.8	192	6	US-10-603-114-7828	Sequence 7828, Ap
55	7	0.8	193	6	US-10-603-114-6794	Sequence 6794, Ap
56	7	0.8	194	6	US-10-293-244-3807	Sequence 3807, Ap
57	7	0.8	231	6	US-10-374-780A-414	Sequence 414, App
58	7	0.8	231	6	US-10-374-780A-1246	Sequence 1246, Ap
59	7	0.8	251	6	US-10-603-114-4250	Sequence 4250, Ap
60	7	0.8	258	6	US-10-293-244-3649	Sequence 3649, Ap
61	7	0.8	269	6	US-10-603-114-5243	Sequence 5243, Ap
62	7	0.8	271	6	US-10-292-798-1056	Sequence 1056, Ap
63	7	0.8	287	6	US-10-603-114-5282	Sequence 5282, Ap
64	7	0.8	295	5	US-09-665-308E-2	Sequence 2, Appl
65	6	0.8	303	6	US-10-374-780A-1658	Sequence 1658, Ap
66	7	0.8	325	6	US-10-610-917-5	Sequence 5, Appl
67	7	0.8	329	6	US-10-292-798-798	Sequence 798, App
68	7	0.8	332	6	US-10-293-244-1681	Sequence 1681, Ap
69	7	0.8	333	1	PCT-US03-21576-38	Sequence 38, Appl
70	7	0.8	333	1	PCT-US03-21576-39	Sequence 39, Appl
71	7	0.8	341	1	PCT-US03-21703-13	Sequence 13, Appl
72	7	0.8	341	1	PCT-US03-21703-13	Sequence 13, Appl
73	7	0.8	345	6	US-10-387-629-106	Sequence 106, App
74	7	0.8	351	6	US-10-374-780A-124	Sequence 124, App
75	7	0.8	352	6	US-10-341-434-178	Sequence 178, App
76	7	0.8	354	7	US-60-487-369-6	Sequence 6, Appl
77	7	0.8	355	6	US-10-403-902A-22	Sequence 22, Appl
78	7	0.8	357	6	US-10-293-244-1353	Sequence 1353, Ap
79	7	0.8	359	6	US-10-603-114-6082	Sequence 6082, Ap
80	7	0.8	360	6	US-10-374-780A-2376	Sequence 2376, Ap
81	7	0.8	370	7	US-60-487-610-1928	Sequence 1928, Ap
82	7	0.8	375	7	US-10-408-765A-1649	Sequence 1649, Ap
83	7	0.8	384	1	PCT-US03-15466-4	Sequence 4, Appl
84	7	0.8	385	6	US-10-603-114-5737	Sequence 5737, Ap

```
85 7 0.8 411 6 US-10-408-765A-126 Sequence 126, App
86 7 0.8 420 5 US-10-273-573-6661 Sequence 6661, Ap
87 7 0.8 425 5 US-09-721-870-109 Sequence 109, App
88 7 0.8 430 6 US-10-369-072-91 Sequence 91, Appl
89 7 0.8 438 6 US-10-282-798-1844 Sequence 1844, Ap
90 7 0.8 450 6 US-10-394-322A-15 Sequence 15, Appl
91 7 0.8 455 6 US-10-273-573-7309 Sequence 7309, Ap
92 7 0.8 456 6 US-10-461-791-10 Sequence 10, Appl
93 7 0.8 536 6 US-10-461-792-10 Sequence 10, Appl
94 7 0.8 557 6 US-10-326-956-439 Sequence 439, App
95 7 0.8 575 6 US-10-617-624-8 Sequence 8, Appl
96 7 0.8 575 6 US-10-620-061-8 Sequence 8, Appl
97 7 0.8 586 6 US-10-408-765A-1266 Sequence 1266, Ap
98 7 0.8 596 6 US-10-273-573-5878 Sequence 5878, Ap
99 7 0.8 617 6 US-10-273-573-5880 Sequence 5880, Ap
100 7 0.8 617 6 US-10-273-573-7306 Sequence 7306, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-357-820-28
; Sequence 28, Application US/10357820
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Casman, Stacie J.;
; APPLICANT: Gorman, Linda;
; APPLICANT: Ji, Weizhen;
; APPLICANT: Kekuda, Ramesh;
; APPLICANT: Li, Li;
; APPLICANT: Padigaru, Muralidhara;
; APPLICANT: Patnirajan, Meera;
; APPLICANT: Pena, Carol E.A.;
; APPLICANT: Shenoy, Suresh G.;
; APPLICANT: Shimkete, Richard A.;
; APPLICANT: Stone, David J.;
; APPLICANT: Taupier, Raymond J.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-538B
; CURRENT APPLICATION NUMBER: US/10/357,820
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/679460
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: 09/730617
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 10/074978
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 10/138588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/387002
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/355099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/375579
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/393265
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 401825
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 60/381666
; PRIOR FILING DATE: 2002-05-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Cuiaseq1.2 version 0.1
; SEQ ID NO 28
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-357-820-28
Alignment Scores: 7.55e-07 Length: 514
Pred. No.: 6
```

```
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-357-820-28 (1-514)

QY 1028 AGAGCTGCTGTAATTAATGACAACTTCTCAAGCTCGTCTG 1072

DB 274 ATGSetleuValAsnlystrPthrPheleuylsAlaArgleu 288

RESULT 2

US-10-408-765A-2003

; Sequence 2003, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465 IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2003

; LENGTH: 636

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-2003

Alignment Scores:

Pred. No.: 7.43e-07 Length: 636

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.68% Indels: 0

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-2003 (1-636)

QY 1028 AGAGCTGCTGTAATTAATGACAACTTCTCAAGCTCGTCTG 1072

DB 164 ATGSetleuValAsnlystrPthrPheleuylsAlaArgleu 178

RESULT 3

PCT-US03-09929-92

; Sequence 92, Application PCT/US0309929

; GENERAL INFORMATION:

; APPLICANT: Caragen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-5738-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

```

; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 92
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-92

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-92 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 4
PCT-US03-09929-94
; Sequence 94, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 94
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-94

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
```

```

DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x PCT-US03-09929-94 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 5
PCT-US03-09929-96
; Sequence 96, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 96
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-96

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-96 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 6
PCT-US03-09929-98
; Sequence 98, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
```

```
/ PRIOR APPLICATION NUMBER: 60/370,381
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: 60/370,969
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: 60/371,002
/ PRIOR FILING DATE: 2002-04-09
/ PRIOR APPLICATION NUMBER: 60/372,002
/ PRIOR FILING DATE: 2002-04-12
/ PRIOR APPLICATION NUMBER: 60/384,297
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/386,816
/ PRIOR FILING DATE: 2002-06-07
/ PRIOR APPLICATION NUMBER: 60/389,123
/ PRIOR FILING DATE: 2002-06-13
/ PRIOR APPLICATION NUMBER: 60/402,207
/ PRIOR FILING DATE: 2002-08-09
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 179
/ SOFTWARE: CuroSeqList version 0.1
/ SEQ ID NO 98
/ LENGTH: 775
/ TYPE: PRT
/ ORGANISM: Homo sapiens
PCT-US03-09929-98

Alignment Scores:
Pred. No.: 7,57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
Gaps: 0
DB: 1

US-09-774-490-1 (1-2709) x PCT-US03-09929-98 (1-775)

QY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTTGAGATGT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaItrpAspGly 541

RESULT 7
/ Sequence 97, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ FILE REFERENCE: 21402-573B-061
/ CURRENT APPLICATION NUMBER: PCT/US03/09929
/ CURRENT FILING DATE: 2003-04-01
/ PRIOR APPLICATION NUMBER: 60/368,996
/ PRIOR FILING DATE: 2002-04-01
/ PRIOR APPLICATION NUMBER: 60/369,980
/ PRIOR FILING DATE: 2002-04-04
/ PRIOR APPLICATION NUMBER: 60/370,381
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: 60/370,969
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: 60/371,002
/ PRIOR FILING DATE: 2002-04-09
/ PRIOR APPLICATION NUMBER: 60/372,002
/ PRIOR FILING DATE: 2002-04-12
/ PRIOR APPLICATION NUMBER: 60/384,297
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/386,816
/ PRIOR FILING DATE: 2002-06-07
/ PRIOR APPLICATION NUMBER: 60/389,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
```

```
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 97
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Domain search
US-10-369-072-97

Alignment Scores:
Pred. No.: 0.00819 Length: 431
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
Gaps: 0
DB: 6

US-09-774-490-1 (1-2709) x US-10-369-072-97 (1-431)

QY 1040 AATAATGACAACTTCTCAAGCTGCTG 1072
DB 221 AaMySTpThrThrPheLeuYsAlaArgLeu 231

RESULT 8
/ Sequence 4, Application PC/TUS0309929
/ GENERAL INFORMATION:
/ APPLICANT: Curagen Corporation, et al.
/ FILE REFERENCE: 21402-573B-061
/ CURRENT APPLICATION NUMBER: PCT/US03/09929
/ CURRENT FILING DATE: 2003-04-01
/ PRIOR APPLICATION NUMBER: 60/368,996
/ PRIOR FILING DATE: 2002-04-01
/ PRIOR APPLICATION NUMBER: 60/369,980
/ PRIOR FILING DATE: 2002-04-04
/ PRIOR APPLICATION NUMBER: 60/370,381
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: 60/370,969
/ PRIOR FILING DATE: 2002-04-08
/ PRIOR APPLICATION NUMBER: 60/371,002
/ PRIOR FILING DATE: 2002-04-09
/ PRIOR APPLICATION NUMBER: 60/372,002
/ PRIOR FILING DATE: 2002-04-12
/ PRIOR APPLICATION NUMBER: 60/384,297
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/386,816
/ PRIOR FILING DATE: 2002-06-07
/ PRIOR APPLICATION NUMBER: 60/389,123
/ PRIOR FILING DATE: 2002-06-13
/ PRIOR APPLICATION NUMBER: 60/402,207
/ PRIOR FILING DATE: 2002-08-09
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 179
/ SOFTWARE: CuroSeqList version 0.1
/ SEQ ID NO 4
/ LENGTH: 495
/ TYPE: PRT
/ ORGANISM: Homo sapiens
```

```
PCT-US03-09929-4
Alignment Scores:
Pred. No.: 0.0825 Length: 495
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-4 (1-495)

Qy 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 232 Aspaaplyvaltyrphhepheargjlu 241

RESULT 9
PCT-US03-09929-2
; Sequence 2, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Curaseq1st version 0.1
; SEQ ID NO 2
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-2

Alignment Scores:
Pred. No.: 0.0809 Length: 634
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-2 (1-634)

Qy 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 212 Aspaaplyvaltyrphhepheargjlu 221

RESULT 10
PCT-US03-09929-8
; Sequence 8, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.

PCT-US03-09929-2
; Sequence 10, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-5738-061
CURRENT APPLICATION NUMBER: PCT/US03/09929
CURRENT FILING DATE: 2003-04-01
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123

Alignment Scores:
Pred. No.: 0.0809 Length: 638
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-8 (1-638)

Qy 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 214 Aspaaplyvaltyrphhepheargjlu 223

RESULT 11
PCT-US03-09929-10
; Sequence 10, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
```


;; PRIOR FILING DATE: 2002-06-13
;; PRIOR APPLICATION NUMBER: 60/402,207
;; PRIOR FILING DATE: 2002-08-09
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; SOFTWARE: Curaseqlist version 0.1
;; SEQ ID NO 10
;; LENGTH: 638
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US03-09929-10

Alignment Scores:
Pred. No.: 0.0809
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-10 (1-638)
QY 917 GATGACAAGTATCTTTCTTCCTCCGTGA 946
DB 214 AspAspIyValIyTrPhepPheArGlu 223

RESULT 12
PCT-US03-09929-6

;; Sequence 6, Application PC/TUS0309929
;; GENERAL INFORMATION:
;; APPLICANT: Curagen Corporation, et al.
;; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;; FILE REFERENCE: 21402-573B-061
;; CURRENT APPLICATION NUMBER: PCT/US03/09929
;; CURRENT FILING DATE: 2003-04-01
;; PRIOR APPLICATION NUMBER: 60/368,996
;; PRIOR FILING DATE: 2002-04-01
;; PRIOR APPLICATION NUMBER: 60/369,980
;; PRIOR FILING DATE: 2002-04-04
;; PRIOR APPLICATION NUMBER: 60/370,381
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: 60/370,969
;; PRIOR FILING DATE: 2002-04-08
;; PRIOR APPLICATION NUMBER: 60/371,002
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/372,002
;; PRIOR FILING DATE: 2002-04-12
;; PRIOR APPLICATION NUMBER: 60/384,297
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: 60/386,816
;; PRIOR FILING DATE: 2002-06-07
;; PRIOR APPLICATION NUMBER: 60/389,123
;; PRIOR FILING DATE: 2002-06-13
;; PRIOR APPLICATION NUMBER: 60/402,207
;; PRIOR FILING DATE: 2002-08-09
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 179
;; SOFTWARE: Curaseqlist version 0.1
;; SEQ ID NO 6
;; LENGTH: 833
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US03-09929-6

Alignment Scores:
Pred. No.: 0.0793
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-6 (1-833)

QY 917 GATGACAAGTATCTTTCTTCCTCCGTGA 946
DB 232 AspAspIyValIyTrPhepPheArGlu 241

RESULT 13
PCT-US03-09929-12

;; Sequence 12, Application PC/TUS0309929
;; GENERAL INFORMATION:
;; APPLICANT: Curagen Corporation, et al.
;; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;; FILE REFERENCE: 21402-573B-061
;; CURRENT APPLICATION NUMBER: PCT/US03/09929
;; CURRENT FILING DATE: 2003-04-01
;; PRIOR APPLICATION NUMBER: 60/368,996
;; PRIOR FILING DATE: 2002-04-01
;; PRIOR APPLICATION NUMBER: 60/369,980
;; PRIOR FILING DATE: 2002-04-04
;; PRIOR APPLICATION NUMBER: 60/370,381
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: 60/370,969
;; PRIOR FILING DATE: 2002-04-08
;; PRIOR APPLICATION NUMBER: 60/371,002
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/372,002
;; PRIOR FILING DATE: 2002-04-12
;; PRIOR APPLICATION NUMBER: 60/384,297
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: 60/386,816
;; PRIOR FILING DATE: 2002-06-07
;; PRIOR APPLICATION NUMBER: 60/389,123
;; PRIOR FILING DATE: 2002-06-13
;; PRIOR APPLICATION NUMBER: 60/402,207
;; PRIOR FILING DATE: 2002-08-09
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 179
;; SOFTWARE: Curaseqlist version 0.1
;; SEQ ID NO 12
;; LENGTH: 833
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US03-09929-12

Alignment Scores:
Pred. No.: 0.0793
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-12 (1-833)

QY 1043 AATGACAACATCTCTCAAGCTGCTG 1072
DB 274 LysTrpThrThrPheLeuLysIaArgIeu 283

RESULT 14
US-10-369-072-14

;; Sequence 14, Application US/10369072
;; GENERAL INFORMATION:
;; APPLICANT: Alsobrook II, John P
;; APPLICANT: Spaderna, Stephen K
;; APPLICANT: Tchernev, Velizar
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Shenoy, Suresh
;; APPLICANT: Spytek, Kimberly
;; APPLICANT: Zernusen, Bryan
;; APPLICANT: Patuturajan, Meera
;; APPLICANT: Taupier, Raymond T
;; APPLICANT: Rastelli, Luca
;; APPLICANT: Grosse, William M
;; APPLICANT: Szerkes, Edward S

```

; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-369-072-14

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-14 (1-833)

QY 1043 AATGACACATCTCTCAAGCTGCTG 1072
Db 274 LysTPTnThrPhelulysAlaArgLeu 283

RESULT 15
; Sequence 16, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchermnev, Vellizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zetnusen, Bryan
; APPLICANT: Patnuzajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Raestelli, Luca
; APPLICANT: Groesse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
```

```

; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-369-072-16

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-16 (1-833)

QY 917 GATGACAAAGTACTTTTCTCGGTGA 946
Db 232 AspaaplysvAltyrPhepneheargGlu 241

RESULT 16
; Sequence 18, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchermnev, Vellizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zetnusen, Bryan
; APPLICANT: Patnuzajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Raestelli, Luca
; APPLICANT: Groesse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
```

```
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-18

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-18 (1-833)

QY 917 GATGACAAAGTATCTTTCTTCCTCCGTGAA 946
Db 232 AspAspLysValTyrPhePhePheArgGlu 241

RESULT 17
US-10-369-072-47
/ Sequence 47, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderma, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patlurajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkeres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
```

```
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 47
/ LENGTH: 834
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-369-072-47

Alignment Scores:
Pred. No.: 0.0793 Length: 834
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-47 (1-834)

QY 1043 AAATGACACATCTCTCAAGCTGCTCG 1072
Db 274 LysTPTThrPheLeuLysAlaArgLeu 283

RESULT 18
US-10-369-072-46
/ Sequence 46, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderma, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patlurajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkeres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
```

```

; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 963
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-46

Alignment Scores:
Pred. No.: 0.0784 Length: 963
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-46 (1-963)

QY 917 GATGACAAAGTACTTTCTTCCTCGTGA 946
DB 362 AepaepLyvalTyPhephehargln 371

RESULT 19
US-10-376-397A-42
; Sequence 42, Application US/10376397A
; GENERAL INFORMATION:
; APPLICANT: Ana L.T.O. NACMENTO et al.
; TITLE OF INVENTION: SURFACE PROTEINS OF LEPTOSPIRA
; FILE REFERENCE: FAPESP 204.1 US (10303399)
; CURRENT APPLICATION NUMBER: US/10/376,397A
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Leptospira interrogans serovar copenhageni
US-10-376-397A-42

Alignment Scores:
Pred. No.: 0.861 Length: 350
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-376-397A-42 (1-350)

QY 798 CTATTTCCGAACCTTTGGGCACAC 824
DB 188 LeuSerSerGluLeuLeuGlyThrThr 196

RESULT 20
US-10-369-072-48
; Sequence 48, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Szytek, Kimberly
; APPLICANT: Szytek, Kimberly
; APPLICANT: Zernusen, Bryan
```

```

; APPLICANT: Paturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-48

Alignment Scores:
Pred. No.: 0.837 Length: 510
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-48 (1-510)

QY 1784 CTCGCCGAGACCTTACTGTGCTTG 1810
DB 188 LeuAlaArgAspProTyxCysAlaTrp 196

RESULT 21
US-10-369-072-39
; Sequence 39, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Szytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
```

```
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkete, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 39
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Mus musculus
/ US-10-369-072-39

Alignment Scores:
Pred. No.: 0 81      Length: 782
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-39 (1-782)

QY 1784 CTCGCCGAGACCTTACTGTGCTTGG 1810
Db 483 LeuAlaArgAspProtyrCysAlaIrrp 491

RESULT 22
US-10-369-072-38
/ Sequence 38, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsbrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Vellizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Splytek, Kimberly
/ APPLICANT: Zehusen, Bryan
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkerez, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkete, Richard
/ APPLICANT: Padigaru, Muralidhara
```

```
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 38
/ LENGTH: 893
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-369-072-38

Alignment Scores:
Pred. No.: 0 802      Length: 893
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-38 (1-893)

QY 1784 CTCGCCGAGACCTTACTGTGCTTGG 1810
Db 593 LeuAlaArgAspProtyrCysAlaIrrp 601

RESULT 23
US-10-326-956-1715
/ Sequence 1715, Application US/10326956
/ GENERAL INFORMATION:
/ APPLICANT: Bauer et al.
/ TITLE OF INVENTION: Protein Complexes and Methods for their Use
/ FILE REFERENCE: 220615
/ CURRENT APPLICATION NUMBER: US/10/326,956
/ PRIOR FILING DATE: 2002-12-20
/ PRIOR APPLICATION NUMBER: EP 01 130 253.6
/ PRIOR FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 3282
/ SOFTWARE: Patent In version 3.1
/ SEQ ID NO 1715
/ LENGTH: 1244
/ TYPE: PRT
/ ORGANISM: Saccharomyces cerevisiae
/ US-10-326-956-1715

Alignment Scores:
Pred. No.: 0 783      Length: 1244
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-326-956-1715 (1-1244)

QY 380 CTTTGATGAGGACGAGTAGGCTG 406
|||||
Db 607 LeuLeuAspGluGluArgSerArgLeu 615

RESULT 24

US-10-408-765A-2436

/ Sequence 2436, Application US/10408765A

/ GENERAL INFORMATION:

/ APPLICANT: Ghosh, Soumitra S.

/ APPLICANT: Fahy, Eoin D.

/ APPLICANT: Zhang, Bing

/ APPLICANT: Gibson, Bradford W.

/ APPLICANT: Taylor, Steven W.

/ APPLICANT: Glenn, Gary M.

/ APPLICANT: Warnock, Dale E.

/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

/ FILE REFERENCE: 660088, 465

/ CURRENT APPLICATION NUMBER: US/10/408,765A

/ PRIORITY FILING DATE: 2003-04-04

/ NUMBER OF SEQ ID NOS: 3077

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 2436

/ LENGTH: 3647

/ TYPE: PRF

/ ORGANISM: Homo sapiens

US-10-408-765A-2436

Alignment Scores:

Pred. No.: 0.721 Length: 3647
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-2436 (1-3647)

QY 67 CGTGGTCGACACCCCTCGACGTGACA 41
|||||

Db 2233 ArgGlySerThrProLeuAspValAla 2241

RESULT 25

US-60-487-610-1645

/ Sequence 1645, Application US/60487610

/ GENERAL INFORMATION:

/ APPLICANT: Cargill, Michele

/ APPLICANT: Huang, Hongjin

/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

/ TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,

/ FILE REFERENCE: CL001469

/ CURRENT APPLICATION NUMBER: US/60/487,610

/ PRIORITY FILING DATE: 2003-07-17

/ NUMBER OF SEQ ID NOS: 97101

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 1645

/ LENGTH: 275

/ TYPE: PRF

/ ORGANISM: Homo sapiens

US-60-487-610-1645

Alignment Scores:

Pred. No.: 8.92 Length: 275
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 7 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-487-610-1645 (1-275)

QY 1780 TTGCTCGCCGAGACCCCTTACTG 1803
|||||
Db 22 LeuProArgProArgProLeuLeu 29

RESULT 26

US-10-292-798-748

/ Sequence 748, Application US/10292798

/ GENERAL INFORMATION:

/ APPLICANT: Sowa, Makiko

/ APPLICANT: Asai, Kiyoshi

/ APPLICANT: Akiyama, Yutaka

/ APPLICANT: Abaratani, Hiroyuki

/ TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS

/ FILE REFERENCE: 08435/166

/ CURRENT APPLICATION NUMBER: US/10/292,798

/ PRIORITY FILING DATE: 2002-11-13

/ PRIOR FILING DATE: 2001-12-18

/ PRIOR APPLICATION NUMBER: JP 2001-246789

/ PRIORITY FILING DATE: 2001-06-18

/ NUMBER OF SEQ ID NOS: 2070

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 748

/ LENGTH: 312

/ TYPE: PRF

/ ORGANISM: Homo sapiens

US-10-292-798-748

Alignment Scores:

Pred. No.: 8.84 Length: 312
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-748 (1-312)

QY 1770 GTGCTGAGTGTGCTCGCCGAG 1793
|||||

Db 225 ValLeuSerValAlaSerProGlu 232

RESULT 27

US-10-369-072-10

/ Sequence 10, Application US/10369072

/ GENERAL INFORMATION:

/ APPLICANT: Alsobrook II, John P

/ APPLICANT: Spaderna, Stephen K

/ APPLICANT: Tchernev, Velizar

/ APPLICANT: Liu, Xiaohong

/ APPLICANT: Shenoy, Suresh

/ APPLICANT: Spytek, Kimberly

/ APPLICANT: Zernusen, Bryan

/ APPLICANT: Patnirajan, Meera

/ APPLICANT: Taupier, Raymond T

/ APPLICANT: Rastelli, Luca

/ APPLICANT: Grosse, William M

/ APPLICANT: Szerkies, Edward S

/ APPLICANT: Lepley, Denise M

/ APPLICANT: Shen, Lei

/ APPLICANT: Burgess, Catherine E

/ APPLICANT: Shinkels, Richard

/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same

/ FILE REFERENCE: 21402-050 CON2

/ CURRENT APPLICATION NUMBER: US/10/369,072

/ PRIORITY FILING DATE: 2003-02-18

/ PRIOR FILING DATE: 2002-06-17

/ PRIOR APPLICATION NUMBER: 09/898,994

/ PRIOR FILING DATE: 2001-07-03

```
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 754
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-10

Alignment Scores:
Pred. No.: 8 27      Length: 754
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-10 (1-754)

QY 1043 AATGGACAACATTCCTCAAGCT 1066
DB 275 LysTrpThrThrPheLeuLysAla 282

RESULT 28
US-10-369-072-8
/ Sequence 8, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchervnev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
```

```
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 756
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-8

Alignment Scores:
Pred. No.: 8 27      Length: 756
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-8 (1-756)

QY 1043 AATGGACAACATTCCTCAAGCT 1066
DB 282 LysTrpThrThrPheLeuLysAla 289

RESULT 29
US-10-369-072-36
/ Sequence 36, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchervnev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
```


PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/218,992
PRIOR FILING DATE: 2000-07-17
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 760
TYPE: PRT
ORGANISM: Mus musculus
US-10-369-072-36

Alignment Scores:
Pred. No.: 8.26 Length: 760
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-36 (1-760)

Qy 578 ACTCACTTGACGCTGTGAACG 601
Db 137 ThrlsleuTyrlAcyselgYthr 144

RESULT 30
US-10-293-244-3947
Sequence 3947, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 3947
LENGTH: 761
TYPE: PRT
ORGANISM: Homo sapiens
US-10-293-244-3947

Alignment Scores:
Pred. No.: 8.26 Length: 761
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-3947 (1-761)
Qy 1043 AATGACACATTCCTCAAGCT 1066
Db 282 LysTpTnTrPhleuTyAla 289

RESULT 31
US-10-369-072-35
Sequence 35, Application US/10369072
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John P
APPLICANT: Spaderna, Stephen K
APPLICANT: Tchernev, Velizar
APPLICANT: Liu, Xiaohong
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Taupier, Raymond T
APPLICANT: Rastelli, Luca
APPLICANT: Grosse, William M
APPLICANT: Szerkeres, Edward S
APPLICANT: Lepley, Denise M
APPLICANT: Shen, Lei
APPLICANT: Burgess, Catherine E
APPLICANT: Shinkels, Richard
APPLICANT: Padigaru, Muralidhara
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-050 CON2
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: 10/174,372
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 09/898,994
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/216,585
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,586
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/218,992
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 762
TYPE: PRT
ORGANISM: Homo sapiens
US-10-369-072-35

Alignment Scores:
Pred. No.: 8.26 Length: 762
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-35 (1-762)
Qy 1043 AATGACACATTCCTCAAGCT 1066
Db 282 LysTpTnTrPhleuTyAla 289

```

RESULT 32
US-10-369-072-49
; Sequence 49, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spylek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patruşajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szeikeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralihara
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-050 CON2
CURRENT APPLICATION NUMBER: US/10/369,072
PRIOR FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: 10/174,372
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 09/898,994
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/216,585
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,586
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/218,992
PRIOR FILING DATE: 2000-07-17
Remaining Prior Application data removed - See File Wrapper or PALM
NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-49

Alignment Scores:
Pred. No.:      8.2      Length:      838
Score:          8.00     Matches:      8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      0.89%   Indels:      0
DB:               6      Gaps:      0

US-09-774-490-1 (1-2709) x US-10-369-072-49 (1-838)
OY      578 ACTGACTGTGACGCTGTGGAGACG 601
      |||||
Db      128 ThrlsleutyralaCySgLyThr 135

RESULT 33
US-10-273-573-9338
; Sequence 9338, Application US/10273573
; GENERAL INFORMATION:

```

```

: APPLICANT: Hysqg, Inc
: TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 21272-066
: CURRENT APPLICATION NUMBER: US/10/273,573
: CURRENT FILING DATE: 2002-10-18
: PRIOR APPLICATION NUMBER: 09/522,929
: PRIOR FILING DATE: 2000-04-18
: PRIOR APPLICATION NUMBER: 09/770,160
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 10994
: SOFTWARE: Custom
: SEQ ID NO 9338
: LENGTH: 1074
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: DOMAIN
: LOCATION: (598)..(626)
: OTHER INFORMATION: PEPCRSOR GLYCOPROTEIN SIGNAL RE domain identified by
: OTHER INFORMATION: eMATRIX, accession number PD01719A, p-value=6.100e-15, raw score=
: OTHER INFORMATION: 12.89
: FEATURE:
: NAME/KEY: DOMAIN
: LOCATION: (58)..(468)
: OTHER INFORMATION: Sema domain identified by Pfam, accession name Sema, E-value=
: OTHER INFORMATION: 4.6e-196, Pfam score of 664.7
: US-10-273-9338
: Alignment Scores:
: Pred. NO.:
: Score: 8.05 Length: 1074
: Percent Similarity: 8.00 Matches: 8
: Best Local Similarity: 100.00% Conservative: 0
: Query Match: 100.00% Mismatches: 0
: DB: 0.89% Indels: 0
: Gaps: 0
: US-09-774-490-1 (1-2709) x US-10-273-9338 (1-1074)
: QY 929 TACTTTTCTTCGCTGAATGCA 952
: |||||
: Db 230 TyrphepnearggluAenAla 237
: |||||
: RESULT 34
: US-10-369-072-41
: Sequence 41, Application US/10369072
: GENERAL INFORMATION:
: APPLICANT: Alsebrook II, John P
: APPLICANT: Spaderna, Stephen K
: APPLICANT: Tchenev, Velizar
: APPLICANT: Liu, Xiaohong
: APPLICANT: Shenoy, Suresh
: APPLICANT: Spytek, Kimberly
: APPLICANT: Zerhusen, Bryan
: APPLICANT: Patturajan, Meera
: APPLICANT: Taupier, Raymond T
: APPLICANT: Rastelli, Luca
: APPLICANT: Grose, William M
: APPLICANT: Szerkes, Edward S
: APPLICANT: Lepley, Denise M
: APPLICANT: Shen, Lei
: APPLICANT: Burgess, Catherine E
: APPLICANT: Shimkets, Richard
: APPLICANT: Padigaru, Muralidhara
: TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 21402-050 CON2
: CURRENT APPLICATION NUMBER: US/10/369,072
: CURRENT FILING DATE: 2003-02-18
: PRIOR APPLICATION NUMBER: 10/174,372
: PRIOR FILING DATE: 2002-06-17
: PRIOR APPLICATION NUMBER: 09/898,994
: PRIOR FILING DATE: 2001-07-03
: PRIOR APPLICATION NUMBER: 60/215,854
: PRIOR FILING DATE: 2000-07-03

```

```
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 41
/ LENGTH: 1074
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-369-072-41

Alignment Scores:
Pred. No.: 8.05 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-41 (1-1074)

QY 929 TACTTTTCTCGGAAATGCA 952
DB 230 TYRPNPNEPNEARGLUSMAla 237

RESULT 35
US-10-369-072-44
/ Sequence 44, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patnrajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkerez, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
```

```
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 44
/ LENGTH: 1074
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-369-072-44

Alignment Scores:
Pred. No.: 8.05 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-44 (1-1074)

QY 929 TACTTTTCTCGGAAATGCA 952
DB 230 TYRPNPNEPNEARGLUSMAla 237

RESULT 36
US-10-369-072-43
/ Sequence 43, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernusen, Bryan
/ APPLICANT: Patnrajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkerez, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
```

```
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 43
/ LENGTH: 1077
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-369-072-43

Alignment Scores:
Pred. No.: 8.05 Length: 1077
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-43 (1-1077)
QY 929 TACTTTTCTTCGTGAATGCA 952
DB 230 Typphephearglubenala 237

RESULT 37
US-10-408-765A-26
/ Sequence 26, Application US/10408765A
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Bojn D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Bradford W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
/ APPLICANT: Warnock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
/ FILE REFERENCE: 660088,465
/ CURRENT APPLICATION NUMBER: US/10/408,765A
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3077
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 5032
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-408-765A-26

Alignment Scores:
Pred. No.: 7.16 Length: 5032
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-26 (1-5032)
QY 64 GGGTCACACCCCTCGACGTGCA 41
DB 2266 GlySerThrProLeuAspValAla 2275

RESULT 38
US-10-462-850-2385
/ Sequence 2385, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2399
/ LENGTH: 10
```

```
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2385
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 82-91 and may interact with Sequence 2
US-10-462-850-2385

Alignment Scores:
Pred. No.: 11.7 Length: 10
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2385 (1-10)
QY 2316 TCATGAGCTCATCAACACC 2336
DB 2 SerCySerSerSerThrThr 8

RESULT 39
US-10-462-850-2387
/ Sequence 2387, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2387
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2387

Alignment Scores:
Pred. No.: 11.7 Length: 10
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2387 (1-10)
QY 2316 TCATGAGCTCATCAACACC 2336
DB 3 SerCySerSerSerThrThr 9

RESULT 40
US-10-462-850-2399
/ Sequence 2399, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2399
/ LENGTH: 10
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
;
; FEATURE:
; OTHER INFORMATION: sequence 100
; OTHER INFORMATION: this patent.
US-10-462-850-2399

```

Alignment Scores:	
Pred. No.:	117
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	0.78%
DB:	6
Length:	1
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) X US-10-462-850-2399 (1-10)

Qy	2316	TCATGCAGCTCATCAACCACC	2336
Db	3	SercyserserserThr	9

```

RESULT 41
US-10-462-850-2401
; Sequence 2401, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2401
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
; US-10-462-850-2401

```

Alignment Scores:		
Pred. No.:	117	10
Score:	7.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	0.78*	Indels:
BB:	6	Gaps:
		0

US-09-774-490-1 (1-2709) x US-10-462-850-2401 (1-10)

Qy	2316	TCATGAGCTCATCAACCAACC	2336
Db	1	SercyserserserThrThr	7

```

RESULT 42
US-10-462-850-2485
: Sequence 2485 Application US/10462850
GENERAL INFORMATION:
APPLICANT: Proteom Ltd
TITLE OF INVENTION: Complementary peptide ligands from the human genome
FILE REFERENCE: Human patent
CURRENT APPLICATION NUMBER: US/10/462,850
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 4203
SOFTWARE: ProPatent version 1.0
SEQ ID NO 2485
LENGTH: 10
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
OTHER INFORMATION: sequence located in C9A at 81-90 and may interact with Sequence 2
OTHER INFORMATION: this patent.

```

US-10-462-850-2485	
Alignment Scores:	
Pred. No.:	117
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	0.788
Ds:	6
Gaps:	0
Length:	10
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) X US-10-462-850-2485 (1-10)

Qy	2316	TCATGCAGCTCATCAACCACC	2336
Db	3	SercyserserserThrThr	9

```

RESULT 43
US-10-462-850-2527
: Sequence 2527, Application US/10462850
: GENERAL INFORMATION:
: APPLICANT: Proteom Ltd
: TITLE OF INVENTION: Complementary peptide ligands from the human genome
: FILE REFERENCE: Human patent
: CURRENT APPLICATION NUMBER: US/10/462,850
: CURRENT FILING DATE: 2003-06-17
: NUMBER OF SEQ ID NOS: 4203
: SOFTWARE: ProPatent version 1.0
: SEQ ID NO 2527
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapiens
: FEATURE:
: OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
: US-10-462-850-2527

```

Alignment Scores:	
Pred. No.:	117
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	0.79%
DB:	6
Length:	10
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) X US-10-462-850-2527 (1-10)

Qy	2316	TCATGACGCTCATCAACCACC	2336
Db	1	SercySerserSerserThrThr	7

```

RESULT 44
US-10-462-850-2533
: Sequence 2533, Application US/10462850
: GENERAL INFORMATION:
: APPLICANT: Proteom Ltd
: TITLE OF INVENTION: Complementary peptide ligands from the human genome
: FILE REFERENCE: Human patent
: CURRENT APPLICATION NUMBER: US/10/462,850
: CURRENT FILING DATE: 2003-06-17
: NUMBER OF SEQ. ID NOS: 4203
: SOFTWARE: Protpatent version 1.0
: SEQ ID NO 2533
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapiens
: FEATURE:
: OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
: US-10-462-850-2533

Alignment Scores:
Pred. No.: 117
Score: 7.00
Length: 10
Matches: 7

```

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Conservative: 0
Matches: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2533 (1-10)

QY 2316 TCATGAGCTCATCAACACC 2336
|||||
1 SerCysSerSerSerThrThr 7

RESULT 45

US-10-462-850-2543
Sequence 2543, Application US/10462850
GENERAL INFORMATION:
APPLICANT: Proteom Ltd
TITLE OF INVENTION: Complementary peptide ligands from the human genome
FILE REFERENCE: Human patent
CURRENT APPLICATION NUMBER: US/10/462,850
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 4203
SOFTWARE: Protpatent version 1.0
SEQ ID NO 2543
LENGTH: 10
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2543

Alignment Scores:

Pred. No.: 117
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Length: 10
Matches: 7
Conservative: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2543 (1-10)

QY 2316 TCATGAGCTCATCAACACC 2336
|||||
3 SerCysSerSerSerThrThr 9

RESULT 46

US-10-462-850-2557
Sequence 2557, Application US/10462850
GENERAL INFORMATION:
APPLICANT: Proteom Ltd
TITLE OF INVENTION: Complementary peptide ligands from the human genome
FILE REFERENCE: Human patent
CURRENT APPLICATION NUMBER: US/10/462,850
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 4203
SOFTWARE: Protpatent version 1.0
SEQ ID NO 2557
LENGTH: 10
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-10-462-850-2557

Alignment Scores:

Pred. No.: 117
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Length: 10
Matches: 7
Conservative: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2557 (1-10)

QY 2316 TCATGAGCTCATCAACACC 2336
|||||
1 SerCysSerSerSerThrThr 7

RESULT 47

US-10-601-837-156
Sequence 156, Application US/10601837
GENERAL INFORMATION:
APPLICANT: Holt, Gordon D
APPLICANT: Kelly, Michael D
APPLICANT: Kennedy, Sandra J
APPLICANT: Moyes, Christopher
TITLE OF INVENTION: Proteins, Genes and Their Use for Diagnosis and Treatment of Kidn
TITLE OF INVENTION: Response
FILE REFERENCE: 2543-1-030
CURRENT APPLICATION NUMBER: US/10/601,837
CURRENT FILING DATE: 2003-06-23
PRIOR APPLICATION NUMBER: PCT/GB01/05777
PRIOR FILING DATE: 2001-12-24
PRIOR APPLICATION NUMBER: US 60/260392
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 272
SOFTWARE: PatentIn version 3.1
SEQ ID NO 156
LENGTH: 12
TYPE: PRT
ORGANISM: Rattus Norvegicus
US-10-601-837-156

Alignment Scores:

Pred. No.: 115
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 6
Length: 12
Matches: 7
Conservative: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-601-837-156 (1-12)

QY 1164 GCATTTTGAAGATCTTTAAG 1144
|||||
6 GlyPheLeuGlySerLeuLys 12

RESULT 48

PCT-US02-39397-69
Sequence 69, Application PC/TUS0239397
GENERAL INFORMATION:
APPLICANT: Iowa State University Research Foundation, Inc.
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: RETROELEMENTS
FILE REFERENCE: 08411-031M01
CURRENT APPLICATION NUMBER: PCT/US02/39397
CURRENT FILING DATE: 2003-06-10
PRIOR APPLICATION NUMBER: US 60/339,060
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 168
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 69
LENGTH: 22
TYPE: PRT
ORGANISM: Arabidopsis thaliana
PCT-US02-39397-69

Alignment Scores:

Pred. No.: 110
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 1
Length: 22
Matches: 7
Conservative: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US02-39397-69 (1-22)

OY 2475 CTGTTCTACCTTCTTATTT 2455
|||
DB 16 Leupheleupheleuphe 22

RESULT 49

US-10-369-072-90

/ Sequence 90, Application US/10369072

/ GENERAL INFORMATION:

/ APPLICANT: Alsobrook II, John P

/ APPLICANT: Spaderna, Stephen K

/ APPLICANT: Tchernev, Velizar

/ APPLICANT: Liu, Xiaohong

/ APPLICANT: Shenoy, Suresh

/ APPLICANT: Spytek, Kimberly

/ APPLICANT: Zernhusen, Bryan

/ APPLICANT: Patutajan, Meera

/ APPLICANT: Taupier, Raymond T

/ APPLICANT: Rastelli, Luca

/ APPLICANT: Grosse, William M

/ APPLICANT: Szerkeres, Edward S

/ APPLICANT: Lepley, Denise M

/ APPLICANT: Shen, Lei

/ APPLICANT: Burgess, Catherine E

/ APPLICANT: Shinkets, Richard

/ APPLICANT: Padigaru, Muralidhara

/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same

/ FILE REFERENCE: 21402-050 CON2

/ CURRENT FILING DATE: 2003-02-18

/ PRIOR FILING DATE: 10/174,372

/ PRIOR FILING DATE: 2002-06-17

/ PRIOR FILING DATE: 2001-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR FILING DATE: 2000-07-03

OY 1784 CTGCCGAGACCTTACTGT 1804
|||
DB 13 LeuAlaArgAspProTyrCys 19

RESULT 50

US-10-603-114-7994

/ Sequence 7994, Application US/10603114

/ GENERAL INFORMATION:

/ APPLICANT: GARY BRETON

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

/ FILE REFERENCE: 2709.1002-001

/ CURRENT FILING DATE: 2003-06-24

/ PRIOR FILING DATE: 2003-06-24

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR FILING DATE: 2000-04-05


```
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2152
/ LENGTH: 121
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G1929 Paralogous to G1478
US-10-374-780A-2152

Alignment Scores:
Pred. No.: 96.6 Length: 121
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 6

US-09-774-490-1 (1-2709) x US-10-374-780A-2152 (1-121)

QY 2315 AGCTCTGTACGAGACTTCT 2295
DB 20 SerLeuCystrHrargProser 26

RESULT 52
US-10-273-573-7304
/ Sequence 7304, Application US/10273573
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 21272-066
/ CURRENT APPLICATION NUMBER: US/10/273,573
/ PRIOR FILING DATE: 2002-10-18
/ PRIOR APPLICATION NUMBER: 09/522,929
/ PRIOR FILING DATE: 2000-04-18
/ PRIOR APPLICATION NUMBER: 09/770,160
/ PRIOR FILING DATE: 2001-01-26
/ NUMBER OF SEQ ID NOS: 10994
/ SOFTWARE: Custom
/ SEQ ID NO 7304
/ LENGTH: 126
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (109)..(125)
/ OTHER INFORMATION: C4-TYPE STEROID RECEPTOR ZINC FINGER SIGNATURE domain
/ OTHER INFORMATION: identified by eMUTRIX, accession number PR00047B, p-value=1.000e-
/ OTHER INFORMATION: 20, raw score of 7.63
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (110)..(126)
/ OTHER INFORMATION: Zinc finger, C4 type (two domains) domain identified by Pfam,
/ OTHER INFORMATION: accession name zf-C4, E-value=5.1e-09, Pfam score of 34.4
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(126)
/ OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-7304

Alignment Scores:
Pred. No.: 96.3 Length: 126
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
```

```
DB: 6 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-273-573-7304 (1-126)

QY 1024 CCTCCAAAGTCATTCTTGA 1004
DB 62 ProserLySvalleleuA 68

RESULT 53
US-10-374-780A-1248
/ Sequence 1248, Application US/10374780A
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, Bradley K
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Heard, Jacqueline E
/ APPLICANT: Haake, Volker
/ APPLICANT: Creelman, Robert A
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Adam, Luc J
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Keddie, James
/ APPLICANT: Brown, Pierre E
/ APPLICANT: Pilgrim, Marsha L
/ APPLICANT: Dubell III, Arnold T
/ APPLICANT: Pineda, Omaira
/ APPLICANT: Yu, Guo-Liang
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: MB1-0047 CIP
/ CURRENT APPLICATION NUMBER: US/10/374,780A
/ PRIOR FILING DATE: 2003-02-25
/ PRIOR APPLICATION NUMBER: 09/837,944
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 09/934,455
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/336,049
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1248
/ LENGTH: 145
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Orthologous to G912
US-10-374-780A-1248

Alignment Scores:
Pred. No.: 95.3 Length: 145
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 6

US-09-774-490-1 (1-2709) x US-10-374-780A-1248 (1-145)

QY 1615 AGAAGAGTTCGCTGGAAGA 1635
DB 60 ArgArgGlySerAlaGlyArg 66
```

```
RESULT 54
US-10-374-780A-1342
; Sequence 1342, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keadie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1342
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Orthologous to G1047
US-10-374-780A-1342

Alignment Scores:
Pred. No.: 95          Length: 151
Score: 7.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80%     Indels: 0
DB: 6                 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1342 (1-151)
QY 2253 |||||
DB 28 ProserProserSerieu 34

RESULT 55
US-10-603-114-7828
; Sequence 7828, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
```

```
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7828
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-7828

Alignment Scores:
Pred. No.: 93.3       Length: 192
Score: 7.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%     Indels: 0
DB: 6                 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-7828 (1-192)
QY 1403 TTGCAAGAGTCATCCAGCC 1423
DB 103 PhenlaargserHisProala 109

RESULT 56
US-10-603-114-6794
; Sequence 6794, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6794
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-6794

Alignment Scores:
Pred. No.: 93.2       Length: 193
Score: 7.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%     Indels: 0
DB: 6                 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-6794 (1-193)
QY 27 AGCTTAGTATGATGCCAGC 47
DB 131 SerleuValilleaplaatr 137

RESULT 57
US-10-293-244-3807
; Sequence 3807, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseg, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
```

```
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO: 3807
/ LENGTH: 194
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(194)
/ OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-293-244-3807
```

```
Alignment Scores:
Pred. No.: 93.2 Length: 194
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-293-244-3807 (1-194)

QY 2413 TCCTGGCCCTTGCCGCGCTTG 2393

Db 9 Serritroleuphrothleu 15

RESULT 58

US-10-374-780A-414

/ Sequence 414, Application US/10374780A

/ GENERAL INFORMATION:

/ APPLICANT: Sherman, Bradley K

/ APPLICANT: Riechmann, Jose Luis

/ APPLICANT: Jiang, Cai-Zhong

/ APPLICANT: Heard, Jacqueline E

/ APPLICANT: Haake, Volker

/ APPLICANT: Creelman, Robert A

/ APPLICANT: Ratcliffe, Oliver

/ APPLICANT: Adam, Luc J

/ APPLICANT: Reuber, T. Lynne

/ APPLICANT: Keddie, James E

/ APPLICANT: Brown, Pierre E

/ APPLICANT: Pilgrim, Marsha L

/ APPLICANT: Dubell III, Arnold T

/ APPLICANT: Pineda, Omaliza

/ APPLICANT: Yu, Guo-Liang

/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

/ FILE REFERENCE: MBI-0047 CIP

/ CURRENT APPLICATION NUMBER: US/10/374,780A

/ PRIOR FILING DATE: 2003-02-25

/ PRIOR APPLICATION NUMBER: 09/837,944

/ PRIOR FILING DATE: 2001-04-18

/ PRIOR APPLICATION NUMBER: 60/310,847

/ PRIOR FILING DATE: 2001-08-09

/ PRIOR APPLICATION NUMBER: 09/934,455

/ PRIOR FILING DATE: 2001-08-22

/ PRIOR APPLICATION NUMBER: 60/336,049

```
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO: 414
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G2143 (conserved domain in aa coordinates: 128-179)
US-10-374-780A-414
```

```
Alignment Scores:
Pred. No.: 92 Length: 231
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-374-780A-414 (1-231)

QY 2253 CCATCTCATCATCATCTTTA 2233

Db 73 ProserProserSerleu 79

RESULT 59

US-10-374-780A-1246

/ Sequence 1246, Application US/10374780A

/ GENERAL INFORMATION:

/ APPLICANT: Sherman, Bradley K

/ APPLICANT: Riechmann, Jose Luis

/ APPLICANT: Jiang, Cai-Zhong

/ APPLICANT: Heard, Jacqueline E

/ APPLICANT: Haake, Volker

/ APPLICANT: Creelman, Robert A

/ APPLICANT: Ratcliffe, Oliver

/ APPLICANT: Adam, Luc J

/ APPLICANT: Reuber, T. Lynne

/ APPLICANT: Keddie, James E

/ APPLICANT: Brown, Pierre E

/ APPLICANT: Pilgrim, Marsha L

/ APPLICANT: Dubell III, Arnold T

/ APPLICANT: Pineda, Omaliza

/ APPLICANT: Yu, Guo-Liang

/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

/ FILE REFERENCE: MBI-0047 CIP

/ CURRENT APPLICATION NUMBER: US/10/374,780A

/ PRIOR FILING DATE: 2003-02-25

/ PRIOR APPLICATION NUMBER: 09/837,944

/ PRIOR FILING DATE: 2001-04-18

/ PRIOR APPLICATION NUMBER: 60/310,847

/ PRIOR FILING DATE: 2001-08-09

/ PRIOR APPLICATION NUMBER: 09/934,455

/ PRIOR FILING DATE: 2001-08-22

/ PRIOR APPLICATION NUMBER: 60/336,049

/ PRIOR FILING DATE: 2001-11-19

/ PRIOR APPLICATION NUMBER: 60/338,692

/ PRIOR FILING DATE: 2001-12-11

/ PRIOR APPLICATION NUMBER: 10/171,468

/ PRIOR FILING DATE: 2002-06-14

/ PRIOR APPLICATION NUMBER: 10/225,066

/ PRIOR FILING DATE: 2002-08-09

/ PRIOR APPLICATION NUMBER: 10/225,067

/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1246
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Orthologous to G912
US-10-374-780A-1246

Alignment Scores:
Pred. No.: 92 Length: 231
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1246 (1-231)

QY 1615 AGAAGAGTTCTGTGGAAGA 1635
DB 58 ArgArgGlySerAlaGlyArg 64

RESULT 60
US-10-603-114-4250
/ Sequence 4250, Application US/10603114
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT FILING DATE: 2003-06-24
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US/09/543,681
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 4250
/ LENGTH: 251
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-10-603-114-4250

Alignment Scores:
Pred. No.: 91.4 Length: 251
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-4250 (1-251)

QY 1699 TTGTTGCTGCTTAGTGAAG 1679
DB 177 LeuLeuLeuSerGlyLys 183

RESULT 61
US-10-293-244-3649
/ Sequence 3649, Application US/10293244
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Tang, Y. Tom et al
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 21272-029
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: Not yet Assigned

/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 3649
/ LENGTH: 258
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-293-244-3649

Alignment Scores:
Pred. No.: 91.2 Length: 258
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-3649 (1-258)

QY 728 ACAGATCCCTTTATATAGAT 748
DB 142 ThrAlaSerLeuLeuIleAsp 148

RESULT 62
US-10-603-114-5243
/ Sequence 5243, Application US/10603114
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT FILING DATE: 2003-06-24
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US/09/543,681
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 5243
/ LENGTH: 269
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-10-603-114-5243

Alignment Scores:
Pred. No.: 90.9 Length: 269
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-5243 (1-269)

QY 1593 CTAAGAGACTTGTATGATT 1613
DB 188 LeuArgArgLeuGlyMetIle 194

```
RESULT 63
US-10-292-798-1056
; Sequence 1056, Application US/10292798
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEURATANI, HIROUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1056
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1056

Alignment Scores:
Pred. No.: 90.9      Length: 271
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-1056 (1-271)

QY 1160 AATCCAGTTGTATATGAGTGG 1180
DB 240 AsnProValValtyrGlyVal 246

RESULT 64
US-10-603-114-5282
; Sequence 5282, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5282
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-5282

Alignment Scores:
Pred. No.: 90.5      Length: 287
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-5282 (1-287)

QY 1644 TTTTGGGAGACCGACTGCTA 1664
DB 275 PhepGlyAsnArgLeuLeu 281
```

```
RESULT 65
US-09-665-308E-2
; Sequence 2, Application US/09665308E
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Klein, Theodore M.
; APPLICANT: Odell, Joan T.
; APPLICANT: Orozco, Emil M., Jr.
; TITLE OF INVENTION: PLANT CELL CYCLIN GENES
; FILE REFERENCE: B81149 US NA
; CURRENT APPLICATION NUMBER: US/09/665,308E
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/078,735
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/06047
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 2
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Zea mays
US-09-665-308E-2

Alignment Scores:
Pred. No.: 90.3      Length: 295
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 5                  Gaps: 0

US-09-774-490-1 (1-2709) x US-09-665-308E-2 (1-295)

QY 54 GGTGTGACCCAGCGCTCCGG 74
DB 4 GlyValAspProArgValArg 10

RESULT 66
US-10-374-780A-1658
; Sequence 1658, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omeira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MB1-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
```

```

; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1658
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Orthologous to G1499
US-10-374-780A-1658

Alignment Scores:
Pred. No.: 90.1 Length: 303
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1658 (1-303)
QY 2253 CCATCGATCATCATCTT 2233
DB 52 ProserProserSerleu 58

RESULT 67
US-10-610-917-5
; Sequence 5, Application US/10610917
; GENERAL INFORMATION:
; APPLICANT: Olsen et al.
; TITLE OF INVENTION: Human Freezzed-Like Protein
; FILE REFERENCE: PF389C1
; CURRENT APPLICATION NUMBER: US/10/610,917
; CURRENT FILING DATE: 2003-07-02
; PRIOR APPLICATION NUMBER: US/09/610,650B
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/132,315
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/055,715
; PRIOR FILING DATE: 1997-08-12
; PRIOR APPLICATION NUMBER: 60/081,438
; PRIOR FILING DATE: 1998-04-10
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-610-917-5

Alignment Scores:
Pred. No.: 89.6 Length: 325
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-610-917-5 (1-325)
QY 386 GATGAGAAAGGATGAGCTG 406
DB 261 AspgluGluArgSerArgleu 267

RESULT 68
US-10-292-798-798
; Sequence 798, Application US/10292798
```

```

; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIRO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEHARANT, HIROYUKI
; TITLE OF INVENTION: GONADINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 798
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-798

Alignment Scores:
Pred. No.: 89.5 Length: 329
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-798 (1-329)
QY 1160 AATCAGTGTATGAGTG 1180
DB 298 AsnProValValItyGlyVal 304

RESULT 69
US-10-293-244-1681
; Sequence 1681, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1681
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1681

Alignment Scores:
Pred. No.: 89.5 Length: 332
```

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-1681 (1-332)

QY 728 ACAGCATCCCTTTATAGT 748
DB 145 ThrAlaSerLeuLeuLeu 151

RESULT 70

PCT-US03-21576-38
Sequence 38, Application PC/TUS0321576

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

APPLICANT: KABLE, Amy E.

APPLICANT: SWARNARKAR, Anita

APPLICANT: LEE, Soo Yeun

APPLICANT: RICHARDSON, Thomas W.

APPLICANT: ISON, Craig H.

APPLICANT: KHARE, Reena

APPLICANT: JIN, Pei

APPLICANT: FU, Glenn K.

APPLICANT: BECHA, Shanya D.

APPLICANT: WANG, Jonathan T.

APPLICANT: TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: WARREN, Bridget A.

APPLICANT: JIANG, Xin

APPLICANT: JACKSON, Alan A.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: DUGAN, Brendan M.

APPLICANT: HAPALIA, April J.A.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: TRAN, Uyen K.

APPLICANT: LINDQUIST, Erika A.

APPLICANT: CHAWLA, Narinder K.

APPLICANT: MARQUIS, Joseph P.

TITLE OF INVENTION: SECRETED PROTEINS

FILE REFERENCE: PF-1476 PCT

CURRENT APPLICATION NUMBER: PCT/US03/21576

PRIOR FILING DATE: 2003-07-10

PRIOR APPLICATION NUMBER: US 60/395,722

PRIOR FILING DATE: 2002-07-11

PRIOR APPLICATION NUMBER: US 60/399,003

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/400,909

PRIOR FILING DATE: 2002-08-01

PRIOR APPLICATION NUMBER: US 60/407,552

PRIOR FILING DATE: 2002-08-30

PRIOR APPLICATION NUMBER: US 60/407,202

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PERL Program

SEO ID NO 38

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 7513709CD1

PCT-US03-21576-38

Alignment Scores:

Pred. No.: 89.5

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

US-09-774-490-1 (1-2709) x PCT-US03-21576-38 (1-333)

QY 722 CTGCTGACAGCATCCCTTTA 742
DB 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 71

PCT-US03-21576-39

Sequence 39, Application PC/TUS0321576

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

APPLICANT: KABLE, Amy E.

APPLICANT: SWARNARKAR, Anita

APPLICANT: LEE, Soo Yeun

APPLICANT: RICHARDSON, Thomas W.

APPLICANT: ISON, Craig H.

APPLICANT: KHARE, Reena

APPLICANT: JIN, Pei

APPLICANT: FU, Glenn K.

APPLICANT: BECHA, Shanya D.

APPLICANT: WANG, Jonathan T.

APPLICANT: TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: WARREN, Bridget A.

APPLICANT: JIANG, Xin

APPLICANT: JACKSON, Alan A.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: DUGAN, Brendan M.

APPLICANT: HAPALIA, April J.A.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: TRAN, Uyen K.

APPLICANT: LINDQUIST, Erika A.

APPLICANT: CHAWLA, Narinder K.

APPLICANT: MARQUIS, Joseph P.

TITLE OF INVENTION: SECRETED PROTEINS

FILE REFERENCE: PF-1476 PCT

CURRENT APPLICATION NUMBER: PCT/US03/21576

PRIOR FILING DATE: 2003-07-10

PRIOR APPLICATION NUMBER: US 60/395,722

PRIOR FILING DATE: 2002-07-11

PRIOR APPLICATION NUMBER: US 60/399,003

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/400,909

PRIOR FILING DATE: 2002-08-01

PRIOR APPLICATION NUMBER: US 60/407,552

PRIOR FILING DATE: 2002-08-30

PRIOR APPLICATION NUMBER: US 60/407,202

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PERL Program

SEO ID NO 39

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 7513712CD1

PCT-US03-21576-39

Alignment Scores:

Pred. No.: 89.5

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

Length: 333
Matches: 7
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-21576-39 (1-333)

QY 722 CTGCTGACAGCATCCCTTTA 742

DB 19 LeuLeuThrAlaSerLeuLeu 25


```
RESULT 72
PCT-US03-21703-13
; Sequence 13, Application PC/TUS0321703
; GENERAL INFORMATION:
; APPLICANT: Rupp, Fabio
; APPLICANT: Wang, Jiantui
; APPLICANT: Zhou, Ping
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods and Materials Relating to Novel Polypeptides and Polynuc
; FILE REFERENCE: HYS-B2 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/21703
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/395,402
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: PCT/US02/22858
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/112,944
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/306,971
; PRIOR FILING DATE: 2001-07-21
; PRIOR APPLICATION NUMBER: 10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 60/418,132
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: PCT/US02/39555
; PRIOR FILING DATE: 2002-12-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US03-21703-13

Alignment Scores:
Pred. No.: 89.3      Length: 341
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%   Indels: 0
DB: 1               Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-21703-13 (1-341)
QY 722 CTGCTGACGAGCATCCCTTTTA 742
Db 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 73
US-10-387-629-106
; Sequence 106, Application US/10387629
; GENERAL INFORMATION:
; APPLICANT: Chemcom S.A.
; APPLICANT: Veithen, Alex
; TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors
; FILE REFERENCE: 9409/2192
; CURRENT APPLICATION NUMBER: US/10/387,629
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```
US-10-387-629-106
Alignment Scores:
Pred. No.: 89.2      Length: 345
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80%   Indels: 0
DB: 6               Gaps: 0

US-09-774-490-1 (1-2709) x US-10-387-629-106 (1-345)
QY 36 TTACTAGCTTGTATACATCG 16
Db 167 LeuLeuSerLeuLeuThrSer 173

RESULT 74
US-10-374-780A-124
; Sequence 124, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Uiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MEI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 124
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G597 (conserved domain in AA coordinates: 97-104,137-144)
US-10-374-780A-124

Alignment Scores:
Pred. No.: 89.1      Length: 351
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%   Indels: 0
```

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-124 (1-351)

QY 82 TTGAGCTCGCGCTGTCTCCCA 102

DB 295 LeuSerSerProValLeuPro 301

RESULT 75

US-10-341-434-178

; Sequence 178, Application US/10341434

; GENERAL INFORMATION:

; APPLICANT: OrGene Technologies

; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes

; FILE REFERENCE: 9U 204 205 R1

; CURRENT APPLICATION NUMBER: US/10/341,434

; CURRENT FILING DATE: 2003-07-18

; PRIOR APPLICATION NUMBER: US 60/348,164

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: US 60/348,119

; PRIOR FILING DATE: 2002-01-15

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 178

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-341-434-178

Alignment Scores:

Pred. No.: 89.1 Length: 352

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.80% Indels: 0

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-341-434-178 (1-352)

QY 672 TCCAGCTTAAAAATATTGCC 652

DB 324 SerSerLeuLysIleLeuSer 330

Search completed: August 3, 2003, 10:56:12
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:32:29 ; Search time 24.5 Seconds
(without alignments)
9356.744 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aaccttcttcttcgatg.....agccttcttcttcctaacacc 2709

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 4231058 residues

Word size: 1

Total number of hits satisfying chosen parameters: 557304

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+n2p model -DEV=xjh
-Q/cgcn2_1/USPRO.spool/US09774490/runat 03082003 102928 4485/app query.fasta.1.2887
-DB=Issued_Patents_AA -OPW=faetan -SUFFIX=oligo.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09774490 @CGN 1 1 30 @runat 03082003 102928 4485 -NCPU=6 -ICPU=3
-NO MAMP -LARGEOUTRY -NEG SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgcn2_6/ptodata/2/iaa/5A.COMB.dep:*
2: /cgcn2_6/ptodata/2/iaa/5B.COMB.dep:*
3: /cgcn2_6/ptodata/2/iaa/6A.COMB.dep:*
4: /cgcn2_6/ptodata/2/iaa/6B.COMB.dep:*
5: /cgcn2_6/ptodata/2/iaa/PTUS.COMB.dep:*
6: /cgcn2_6/ptodata/2/iaa/backfile1.dep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	1	US-08-121-713D-54
2	771	86.1	771	1	US-08-835-268-54
3	771	86.1	771	2	US-09-060-692-54
4	771	86.1	771	3	US-08-833-391-54
5	771	86.1	771	4	US-09-060-610-54
6	771	86.1	771	5	PCT-US94-10151A-54
7	655	73.2	655	4	US-08-556-422A-3
8	434	48.5	477	1	US-08-136-922-2
9	13	1.5	775	4	US-09-308-179B-1
10	10	1.1	606	3	US-09-041-236-4
11	10	1.1	606	4	US-09-771-467C-4
12	9	1.0	112	4	US-09-341-461-34

13	9	1.0	295	4	US-08-556-422A-6	Sequence 6, Appl1
14	9	1.0	425	4	US-08-556-422A-7	Sequence 7, Appl1
15	9	1.0	607	4	US-08-556-422A-4	Sequence 4, Appl1
16	9	1.0	630	4	US-09-252-991A-32350	Sequence 32350, A
17	9	1.0	862	4	US-08-556-422A-2	Sequence 2, Appl1
18	9	1.0	913	4	US-08-971-089-4	Sequence 4, Appl1
19	8	0.9	57	4	US-09-152-060-58	Sequence 58, Appl1
20	8	0.9	139	4	US-09-252-991A-24681	Sequence 24681, A
21	8	0.9	332	4	US-09-252-991A-32158	Sequence 32158, A
22	8	0.9	360	4	US-09-252-991A-20193	Sequence 20193, A
23	8	0.9	374	1	US-08-464-148-2	Sequence 2, Appl1
24	8	0.9	374	1	US-08-385-500-2	Sequence 2, Appl1
25	8	0.9	374	1	US-08-846-784-2	Sequence 2, Appl1
26	8	0.9	454	3	US-09-080-044-4	Sequence 4, Appl1
27	8	0.9	454	4	US-09-531-857A-4	Sequence 4, Appl1
28	8	0.9	673	4	US-09-252-991A-25962	Sequence 25962, A
29	8	0.9	849	4	US-09-157-257-4	Sequence 4, Appl1
30	7	0.8	17	1	US-08-644-456-1	Sequence 1, Appl1
31	7	0.8	28	1	US-08-164-151-22	Sequence 22, Appl1
32	7	0.8	37	4	US-09-574-377-27	Sequence 27, Appl1
33	7	0.8	43	4	US-09-205-258-296	Sequence 296, App
34	7	0.8	46	4	US-09-369-247-82	Sequence 82, App
35	7	0.8	48	4	US-09-489-847-175	Sequence 175, App
36	7	0.8	62	4	US-09-328-352-7261	Sequence 7261, Ap
37	7	0.8	62	6	5320958-24	Patent No. 5320958
38	7	0.8	67	4	US-09-107-532A-4497	Sequence 4497, Ap
39	7	0.8	67	4	US-08-456-647B-31	Sequence 31, Appl1
40	7	0.8	68	2	US-08-237-401A-31	Sequence 31, Appl1
41	7	0.8	63	4	US-09-134-001C-5401	Sequence 5401, Ap
42	7	0.8	73	4	US-09-107-532A-5492	Sequence 5492, Ap
43	7	0.8	77	2	US-08-469-537A-47	Sequence 47, Appl1
44	7	0.8	81	2	US-09-014-969-32	Sequence 48, Appl1
45	7	0.8	88	4	US-09-732-210-270	Sequence 270, App
46	7	0.8	90	4	US-09-252-991A-19013	Sequence 19013, A
47	7	0.8	92	4	US-09-370-838-211	Sequence 211, App
48	7	0.8	98	4	US-09-252-991A-28763	Sequence 28763, A
49	7	0.8	101	4	US-09-107-532A-5792	Sequence 5792, Ap
50	7	0.8	101	4	US-08-905-223-437	Sequence 437, App
51	7	0.8	111	4	US-09-107-532A-4380	Sequence 4380, App
52	7	0.8	119	4	US-09-107-532A-3724	Sequence 3724, App
53	7	0.8	119	4	US-09-205-258-268	Sequence 268, App
54	7	0.8	131	4	US-09-107-532A-5741	Sequence 5741, Ap
55	7	0.8	137	4	US-09-252-991A-25510	Sequence 25510, A
56	7	0.8	139	3	US-09-120-663-4	Sequence 4, Appl1
57	7	0.8	133	6	516835-18	Patent No. 516835
58	7	0.8	148	4	US-09-252-991A-18336	Sequence 18336, A
59	7	0.8	150	4	US-09-252-991A-31728	Sequence 31728, A
60	7	0.8	158	3	US-09-053-197A-19	Sequence 19, Appl1
61	7	0.8	158	4	US-09-085-761A-15	Sequence 19, Appl1
62	7	0.8	159	4	US-09-134-001C-5575	Sequence 5575, Ap
63	7	0.8	161	4	US-09-252-991A-31464	Sequence 31464, A
64	7	0.8	162	4	US-09-252-991A-20901	Sequence 20901, A
65	7	0.8	165	4	US-08-311-731A-279	Sequence 279, App
66	7	0.8	166	4	US-09-252-991A-19472	Sequence 19472, A
67	7	0.8	172	2	US-08-469-537A-86	Sequence 86, Appl1
68	7	0.8	172	4	US-09-252-991A-18791	Sequence 18791, A
69	7	0.8	177	4	US-09-252-991A-24167	Sequence 24167, A
70	7	0.8	184	3	US-09-335-409-14	Sequence 14, Appl1
71	7	0.8	184	4	US-09-568-102-14	Sequence 14, Appl1
72	7	0.8	184	4	US-09-567-969-14	Sequence 14, Appl1
73	7	0.8	184	4	US-09-568-480-14	Sequence 14, Appl1
74	7	0.8	184	4	US-09-568-486-14	Sequence 14, Appl1
75	7	0.8	184	4	US-09-568-472-14	Sequence 14, Appl1
76	7	0.8	184	4	US-09-567-899-14	Sequence 14, Appl1
77	7	0.8	184	4	US-09-328-352-6771	Sequence 6771, Ap
78	7	0.8	191	4	US-09-252-991A-16626	Sequence 16626, A
79	7	0.8	192	4	US-09-199-637A-9	Sequence 9, Appl1
80	7	0.8	192	4	US-09-252-991A-25562	Sequence 25562, A
81	7	0.8	198	4	US-09-252-991A-18742	Sequence 18742, A
82	7	0.8	200	4	US-09-252-991A-31710	Sequence 31710, A
83	7	0.8	208	4	US-09-252-991A-27795	Sequence 27795, A
84	7	0.8	214	4	US-09-252-991A-27795	Sequence 27795, A
85	7	0.8	214	4	US-09-252-991A-27795	Sequence 27795, A

86 7 0.8 216 4 US-09-252-991A-22665
 c 87 7 0.8 218 4 US-09-198-452A-621
 88 7 0.8 230 4 US-09-252-991A-30553
 c 89 7 0.8 232 4 US-09-134-001C-5044
 90 7 0.8 234 4 US-09-364-230-10
 91 7 0.8 234 4 US-09-199-637A-61
 c 92 7 0.8 234 4 US-09-107-532A-4817
 93 7 0.8 236 4 US-09-252-991A-26384
 c 94 7 0.8 238 3 US-09-111-470-8
 95 7 0.8 248 6 5169835-15
 96 7 0.8 258 3 US-09-035-706-3
 97 7 0.8 258 3 US-08-955-841-3
 98 7 0.8 258 4 US-09-390-425-3
 99 7 0.8 258 4 US-09-566-906-3
 c 100 7 0.8 261 4 US-09-252-991A-17370

ALIGNMENTS

RESULT 1

US-08-121-713D-54

Sequence 54, Application US/08121713D
 Patent No. 5639856

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.
 APPLICANT: Kolodkin, Alex L.
 APPLICANT: Mathes, David
 APPLICANT: Bentley, David R.
 APPLICANT: O'Connor, Timothy
 TITLE OF INVENTION: The Semaphorin Gene Family
 NUMBER OF SEQUENCES: 100
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 Bush Street, Suite 3200
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/121,713D
 FILING DATE: 13-SEP-1993
 CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: B94-002-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415)343-4341
 TELEFAX: (415) 343-4342
 TELEX:

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:
 LENGTH: 771 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-121-713D-54

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-121-713D-54 (1-771)

200 ATGGCGTGTAAATCAAGATTTCTCTTTCTGGGAGATTAATTAAGACAGACGA 259
 Db MetcilytrpenthraagilevalCysleuphnetpilyalleuleutrralaargala 20
 260 AACATATCAAGATGGGAAGAACATATGTCACAGGCTGAATTTCTTCAAGAAATGTTG 319
 Db AsnlytginbsngilybsasnasvalProargleuylsleuertyrlysglmetleu 40
 21 AsnlytginbsngilybsasnasvalProargleuylsleuertyrlysglmetleu 40
 320 GAATCAACAAATGATGATCACTTTCAATGGGCTGGCCAAAGCTCCACTTTTCATACCTTC 379
 Db Gluserasnasnyallietrpheasngilyleuallasnsersertryhsitphne 60
 41 Gluserasnasnyallietrpheasngilyleuallasnsersertryhsitphne 60
 380 CTTTGGATGAGAACAGGAGTAGGCTGTATGTTGACCAAGATCAATATTTTCATTC 439
 Db LeuenuapglugluaragsersertryvalglylalybsaphsitlepneserPhe 80
 61 LeuenuapglugluaragsersertryvalglylalybsaphsitlepneserPhe 80
 440 GACCTGGTAAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATTTACACAGACGA 499
 Db AspleuvalasnilleybsaphegnilyeilevaltrProvalsertrytrhargarg 100
 81 AspleuvalasnilleybsaphegnilyeilevaltrProvalsertrytrhargarg 100
 500 GATGAATGCAAGTGGGCTGGAAAAAGACATCTGTAAGATGTGCTAATTTTCATCAAGTA 559
 Db Aspelyucyslystrpalaagilybsapleuulysleucysalaaasnpheileysval 120
 101 Aspelyucyslystrpalaagilybsapleuulysleucysalaaasnpheileysval 120
 560 CTTAAGCATATATATCAAGTCACTTGTAGCCCTGTGAACGGGGCTTTTCAATTCAT 619
 Db LeuylsalytrhenginttrhsileuylalyCyselythrghlyalaphenlsproile 140
 121 LeuylsalytrhenginttrhsileuylalyCyselythrghlyalaphenlsproile 140
 620 TGACCTACATTTGAATTTGACATCTCTGAGACAAATTTTAAAGCTGGAGAACTCA 679
 Db Cysrthlyrllleguilllegilyhsishsiprogluabpansnillephelysleugluasnsr 160
 141 Cysrthlyrllleguilllegilyhsishsiprogluabpansnillephelysleugluasnsr 160
 680 CATTGTAAGAAAGCGCGTGGGAAGAGTCATATGACCTTAAGCTGAGACAGATCCCT 739
 Db HispheeluabngilyargilylybserprolyraspProlyseuenuitralaserleu 180
 161 HispheeluabngilyargilylybserprolyraspProlyseuenuitralaserleu 180
 740 TTAATAGATGAGATATATATCTGTGAAGTCAAGCTGATTTTATGGGCGAGACTTGTCT 799
 Db Leuileaspglylguiseuylrserglythralsalasphehetglyargaphehala 200
 181 Leuileaspglylguiseuylrserglythralsalasphehetglyargaphehala 200
 800 ATCTTCCGAATCTTTGGGCAACACACCCATATGACAGACAGACAGATGATTCAGGTGG 859
 Db Ilepheargthrleuengilyhsishsiprogluabpansnillephelysleugluasnsr 220
 201 Ilepheargthrleuengilyhsishsiprogluabpansnillephelysleugluasnsr 220
 860 CTCATATGTCAAAGTTCATTAAGGCCCACTCATCTCAGAGAGTGAAGTCAATCTCGAAGAT 919
 Db LeuasnaspProlyspheilleseeralahleuileseerluserasnasnpheileysval 240
 221 LeuasnaspProlyspheilleseeralahleuileseerluserasnasnpheileysval 240
 920 GACAAAGTATATCTTTCTTCGTAAGATGACATAGATGAGAAACATCTGGAAGAGCT 979
 Db AspelyvaltyrPhePhePheargglyuasnalaileaspilylwhisserserylyvala 260
 241 AspelyvaltyrPhePhePheargglyuasnalaileaspilylwhisserserylyvala 260
 980 ACTCAGCTTGATAGTCAATGATGATGACATGACAAATGACTTTGGAGGGCAGAAAGTCTGGT 1039
 Db ThrtshalaargglyleuglylnileCyslyasnasnpheileysvalasgserserleval 280
 261 ThrtshalaargglyleuglylnileCyslyasnasnpheileysvalasgserserleval 280
 1040 AATAAATGACAACATCTCTCAAGAGCTCGTGAATTTGTCAGTGGCCAGGTCCAAAATGGC 1099
 Db AsnlystrpthrtrpheleuylvalargleuileCyserservalProglyProasngly 300
 281 AsnlystrpthrtrpheleuylvalargleuileCyserservalProglyProasngly 300
 1100 ATTGACACTGATTTTGAAGTCAAGTCAAGATGATATCTTAATGAATTTAAAGATCTCTMAA 1159
 Db IleaspthrhisphesapgluleuglinspvalPheleuuekasnpheylbaspprolys 320
 301 IleaspthrhisphesapgluleuglinspvalPheleuuekasnpheylbaspprolys 320
 1160 AATCCAGTGTATATGAGAGTGTTTAGACTTCCAGTAACATTTTCAAGGAGATCAGCCGTG 1219
 Db AspProvalvaltyrglyvalPheitrThserserasnillepelysglyseralaval 340
 321 AspProvalvaltyrglyvalPheitrThserserasnillepelysglyseralaval 340
 1220 TGTATGATATGAGATGAGTATGAGAAAGGAGTCTCTGTCATATATGCCCACAGGAGAT 1279
 Db CyswelytrisermetseraspsvalArgargvalPheleuulypProlytralashsargasp 360
 341 CyswelytrisermetseraspsvalArgargvalPheleuulypProlytralashsargasp 360
 1280 GGACCCCAATATCAATGGGTGCTTATCAAGAAAGATCCCTATATCAACGGCCAGAACT 1339

Db	361	GlYPProAenTyGIntPrVaIProTyrgInGluValProTyPProAlGProLylYhr	380
QY	1340	TGTCCAGCAAAACATTTGGTGGTTTGACTTCAAAAAGACCTTCTGATGATTTATA	1399
Db	381	CysProSerLysThrPheGlyGlyIYpHeApsSerThrLysAspLeuProAspAraValIle	400
QY	1400	ACCTTGGAAAGACATCCAGCCAGCCATGTAACAATCCAGGTGTTCCATTAACAACACGCCCA	1455
Db	401	ThrPheAlaAspSerHisProAlaMetCylTranProValPheProMetAspAsnArgPro	420
QY	1460	ATAGTATCAAAACCGAGTGAATTAATTAATCAATTTATCAAAATTTGTGTGAGCCAGTGGAT	1519
Db	421	IleValIleLysThrAspValAsnTyrgInPheThrGlnIleValValAspArgValAsp	440
QY	1520	GCAGAAATGGACAGTATGATGTTATGTTTATCGSACAAGATGTTGGACCGTGTCTTAA	1579
Db	441	AlaGluAspGlyGlnTyLAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
QY	1580	GTAGTTCAATTCCTTAAGAGACCTTGATGATTTTAAGAAGAGTTCTGCTGAAGAATA	1639
Db	461	ValValSerIleProLysGlnThrTrpTyLAspLeuGluGluValLeuLeuGlnGlnMet	480
QY	1640	ACAGTTTTCGGGAAACCGACGTCTATTTCAAGATAGAGCTTCCACTTAAGACAGCAAA	1699
Db	481	ThrValPheAArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
QY	1700	CTATATATTTGGTTCAACGGCTGGGGTTGCCCACTCCCTTACAACGGGTGATATTTAC	1759
Db	501	LeuTyrlIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy	520
QY	1760	GGGAAAGCTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTGGAGTGGTCT	1819
Db	521	GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGlySer	540
QY	1820	GCATGTTCTCCGTATTTTCCACTGSCAAAGAAGACCAAGAAGCGCAAGATATAGAAT	1879
Db	541	AlaCysSerAlaGlyTyPheProThrAlaLysArgAlaGlnArgGlnAspIleArgAsn	560
QY	1880	GGAGACCCAGTCACTCACTGTTCAGACTTACAACATGATTAATCAACATGAGCCAGACCT	1939
Db	561	GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisAspGlnHisIleSerPro	580
QY	1940	GAAGAGAGAAATCATTTATGCTGTAGAKAAATGTAGACATTTTGGATGCAATGCCAAG	1999
Db	581	GluGluAspGlyIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProLys	600
QY	2000	TCGCAGAGAGCGCTGATTTATGGCAATTCAGAGGCGCAAAATGAAGAAGCGAAAGAG	2059
Db	601	SerGlnAspAlaLeuValTyLTrpGlnPheGlnArgAspAsnGluGluAspGlnGln	620
QY	2060	ATCAGAGTGAATGATCATATATATAGACAGACATCAAGGCTTGTGCTACATAGTCAAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuHisArgSerLeuGln	640
QY	2120	CAGAAAGATTCAAGCAATTACCTCTGCCATGGCGTGGAACATGGGTTCAACAACCTT	2179
Db	641	GlnLysAspSerGlyAsnTyLLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu	660
QY	2180	CTTAAGGTAAACCTCGAAGTCAATTTAGACACAGAGCATTTGGAAAGACTTCTTCATTAAGAT	2239
Db	661	LeuLysValThrLeuGlnValIleLeuSerThrGlnHisLeuGluGluLeuHisIleLysAsp	680
QY	2240	GATGATGAGATGGCTTAAAGCAAAAGAAATGTCTCATATAGCATAGACACTTAAGCAGAG	2299
Db	681	AspAspGlyAspGlySerLysThrLysGlnMetSerHisSerMetThrProSerGlnLys	700
QY	2300	GTCGTGTCAGAGACTTCATCAAGCTCATCAACACCCCAATTTCAACAGATGATGAG	2359
Db	701	ValTrpTyLArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
QY	2360	TTCTGTGAACAAGTTTGGAAAGGACCGGAAACAACGTGGCAAAAGCGCCAGCATATACC	2419

```

Db      721  PheylsclucInValIrrpYsArYgsbArYgSLInA-gArGInArGProGlyHiEthr 761
Qy      2420  CCAGGAAACGATACAAATGGAAGCACTTACCAAGAAATTAAGAAAGTAGAAACAGAG 2471
Db      741  ProGlyAmsbArSAnSlyStrpYshIslEauGInGInAenLySLYArGArSAnArGArY 761
Qy      2480  ACCACGAATTGAGAGGACCCAGAGACTGTC 2512
Db      761  ThrHlGInpHegInArGAlaProArGArSArVal 771

RESULT 2
US-08-835-268-54
: Sequence 54, Application US/08835268
: Patent No. 5807826
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Matthes, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphorin Gene Family
: NUMBER OF SEQUENCES: 100
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/835,268
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/121,713
: FILING DATE: 13-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)343-4341
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 54:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-835-268-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-835-268-54 (1-771)
Qy      200  ATGGGCTGGTTACTAGAGATTGTCTGTTCTGGGAGATTACTTACGACAGAGCA 259
Db      1  MetGlyTrrpLeuThrArGlgIleValCysLeuPhetrrpGlyValLeuIutrrAlaArGAla 20
Qy      260  AACTATCAGAAATGGGAAGAACATGTGCCAAGCTGAATTAATCTCTCAAGAAGATTGG 319
Db      1  MetGlyTrrpLeuThrArGlgIleValCysLeuPhetrrpGlyValLeuIutrrAlaArGAla 20

```

Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
Qy 320 GAATCAACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
Qy 380 CTTTGGATGAGAACGGAGTAGGCTGATGTGGAGCAAGATCATATTTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
Qy 440 GACCTGGTTATATCAAGAGATTTTCAAAAGATGTGTGGCCAGATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArgArg 100
Qy 500 GATAAAGCAAGTGGGCTGAAAAAGACATCTGAAAGATGGTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysValaAsnPheIleLysVal 120
Qy 560 CTTAAGGCATATTAATCAAGACTCACTGTGACGCTGTGGAAACGGGGCTTTTCATCCAA 619
Db 121 LeuLysAlaTyrAsnGlnThrHisIleLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
Qy 620 TGCACTCAATGATGAATTTGACATCATCTGAGACAAATATTTTAACTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisIleAspGlnAspAsnIlePheLysLeuGluAsnSer 160
Qy 680 CATTTGAAAAACGGCCGTGGGAAGATCCATATGACCTTAAGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuThrAlaSerLeu 180
Qy 740 TTAATAGATGAGAAATTAATCTGTGAACCTGACAGCTGATTTTATGGGCGAGACTTGTCT 799
Db 181 LeuIleAspGlyLeuLysTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCCGAACCTTTGGGACACCAACCCCAATCGAGACAGCAGCATGATCCAGGTGC 859
Db 201 IlePheArgThrLeuGlnLysHisIleAspGlnIleArgThrGlnGlnHisAspSerArgTyr 220
Qy 860 CTCATATGATCCAAAGTTCAATTAGTGGCCACTCATCTCAGAGAGTGAACAATCCGAAAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
Qy 920 GACAAGATATATCTTTTCTCCGTGAATAATGCAATAGTAGAGAACTCTGAAAAAGCT 979
Db 241 AspLysValTyrPhePhePheAsnGluAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
Qy 980 ACTCAACCTGAAATAGCTCAGATATGCAAGAAATGACTTTGGAGGGCAACAAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlnHisArgSerLeuVal 280
Qy 1040 AATAAAGTGAACAATTCCTCAAGCTCGTGTGATTTGCTCAGTCCAGGTCGAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
Qy 1100 ATTGACATCATTTTGAATGAACTGACAGATGTATTTCTTAATGAATTTAAAGATCTTAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy 1160 AATCAAGTGTATATGAGTGTTTTACAATTCCTCAAGTAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AspProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
Qy 1220 TGTATGATAGCATGATGATGATGAGAAAGGTGTCTGTGGTCAATAGCCCAAGGAT 1279
Db 341 CysMetLysSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisAspAsp 360
Qy 1280 GGAACCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTTATCCACGGCCAGAACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
Qy 1340 TGTCCAGCAAAACATTTGGTGTGTTTGAAGTCTTCAAAAGACTTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

Qy 1400 ACCTTGCAAGAGATCATCCAGCATGTACAATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy 1460 ATAGTATCAAAACGAGATGAAATTAATTAATTAACAATAATGTCTGTAACGAGTGGAT 1519
Db 421 IleValIleLysThrAspValaLeuTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy 1520 GCAGAAAGTGAACAGATGATGATGTTATGTTATGGAACAGATGGTGGACCGTCTTAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValaMetPheIleGlyThrAspValaGlyThrValaLeuLys 460
Qy 1580 GTAGTTTCAATTCCTAAGAGACTGTGATGATTTGAAGAGGTTCTGCTGAAGAAATG 1639
Db 461 ValValaSerIleProLysGluThrTrpTyrAspLeuGlnGluValaLeuLeuGluMet 480
Qy 1640 ACAAGTTTTTGGGAACCGACTGCTATTTGACAAATGAGACTTTCATTAACGACAAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
Qy 1700 CTATATATGCTTCAACGGCTGGGGTTCGCCAGCTCCCTTACACGGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValaGlnLeuProLeuHisArgCysAspIleTyr 520
Qy 1760 GGGAAAGCGTGTGAGTGTGCTCGCCCGGAGACCTTACTGTGCTTGGAGTGGTCT 1819
Db 521 GlyLysAlaCysValaGluCysLysLeuAlaAspAspProTyrCysAlaTrpAspGlySer 540
Qy 1820 GCATGTTCTCGTATTTTCCACTGCAAAAGACGCAACAAAGACAGCAAGATTAAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
Qy 1880 GGAACCCACTGACTGACTGCTGAGCTTACACCAATGATTAATGACAGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAsnAspAsnHisIleGlyHisSerPro 580
Qy 1940 GAAGAGAGAATCATCATGATGCTGAGAGATAGTACGACATTTTGGAAATGACAGTCCGAAG 1999
Db 581 GluGluArgGlyIleLeuTyrGlyValaGluAsnSerSerThrPheLeuGluCysSerProLys 600
Qy 2000 TGCAGAGACGCTGTCTATTTGGCAATTCAGAGGCGAAATGAAGACGAAAGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgGlnArgGlnGluArgLysGluGln 620
Qy 2060 ATCAGAGTGAATGATCATCATCAGACAGATCAAGGCTTCTGTAGTGTGATCAACA 2119
Db 621 IleArgValaAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy 2120 CAGAGGATTCAGAGCAATTAACCTGTGCAATGCGGTGGAACATGGTTCATACAACCTTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisIleValaValGlnHisGlyPheIleGlnThrLeu 660
Qy 2180 CTTAAGGTAACCTGGAAGCTGATTAACAACAAGCATTTTGGAGAACTTCTCATTAAGAT 2239
Db 661 LeuLysValaThrLeuGluValaIleAspThrGlnHisIleGluGluLeuLeuHisLysAsp 680
Qy 2240 GATGATGAGATGAGTGGCTTAAAGCCAAAGAAATCCAAATGACATGACACTGACCCGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerLeuThrProSerGlnLys 700
Qy 2300 GTCTGTACAGAGACTTCATGACATGACATCAACACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTrpTyrAspAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
Qy 2360 TTCTGTGAACAAGTTTGAAGAAAGGACGAAACAAAGTGTGGAAAGCCAGGACATACC 2419
Db 721 PheCysGluGlnValaTrpLysArgAspArgGlyGlnArgGlnArgProGlyHisThr 740
Qy 2420 CCAAGGAAACATGAACAATGGAAGCACTTACAAGAAATTAAGAAAGTGAAGAACAGAGAG 2479
Db 741 ProGluAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760

QY 2480 ACCACGAATTGAGAGGACCCAGAGATGTC 2512
Db 761 ThrhsgluphegiuargalaproargserVal 771

RESULT 3

US-09-060-692-54
Sequence 54, Application US/09060692
Patent No. 5935865
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matches, David R.
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,692
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-692-54

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-060-692-54 (1-771)

QY 200 ATGGGCTGTTAACTAGATGTTGCTCTTTCTGGGAGATTAATTACTTACAGCAAGAGCA 259
Db 1 MetGlyTTPeuthrhrarglleValCysleuphetrpolyalleuLeuintrhalaargala 20
QY 260 AACTATCAGATGGAAGAAACAATGTGCCAAGCTGTAATTAATCTTCAAGAAATGTTG 319
Db 21 AsnTyrGlnasnnglylsAsnAsnValProhrgleuylsleuSertYrlysglMetLeu 40
QY 320 GAATCCAAAGATGATCATCTTCAATGGCTGGCCCAAGCTCCAGTTATCATCTTTC 379
Db 41 GluserasnAsnValIlethrPheasnGlyleuAlaasnSertSertYrhlsthrPhe 60

QY 380 CTTTGGATGAGGAAGAGATGAGCTGTATGTTGAGCAAGATCATAATTTTCATTC 439
Db 61 LeuLeuAspGluInuIuargserArgleuYrValGlyAluLysAspHisIlePheSertPhe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGACA 439
Db 81 AspLeuValAsnIleLysAspPheGlnlyslleValItrProValSertYrThrArgArg 100
QY 500 GATGAATGCAAGTGGGCTGGAAAGACATCCCTGAAGAAATGTCCTAATTTTCAACAGTA 559
Db 101 AspGluCysLeuYrtrpAlaGlylsAspIleLeuLysGluCysAlaAsnPheIleVal 120
QY 560 CTTAAGCAATATATCAGACTCCTTGATGCGCTGTGGAACGGGGCTTTTCATCAAT 619
Db 121 LeuLysAlaLetyrasnGlntrhrIstleuYrAlaCysGlyYrthrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAATTTGACATCATCTGAGGACAAATTTTAACTGGCAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisIstProGluAspAsnIlePheLysleuGluAsnSer 160
QY 680 CATTTGAAAGCGCGTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysPyrProYrAspProLysleuLeuIntrhalaSerLeu 180
QY 740 TTAATGATGAGAAATTAATCTCTGGAAGTGCAGCTGATTTTATGGGGCAGACTTGTCT 739
Db 181 LeuIleAspGlyGluLeuYrSerglyThrAlaIleAspPheGlyArgAspPheAla 200
QY 800 ATCTTCGAACTCTTGGGACCAACCAACCAATGAGACAGAGCATGATTCAGAGTGG 859
Db 201 IlePheArgThrLeuGlyHisIstIstProIleArgThrGluGlnHisAspSertArgTyr 220
QY 860 CTCATGATGCAAGTCAATTAAGTGGCCACCTCATCTGAGAGTGAACATCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSertAlaHisIleuIleSertIserAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTTCTCCGTGAAATGCAATGATGAGAAACACTCTGAAAAAGCT 979
Db 241 AspLysValYrPhePhePheArgGluAsnAlaIleAspGlyGlnIstSerglyLysAla 260
QY 980 ACTCAGCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyIstArgSertLeuVal 280
QY 1040 AATTAATGAGCAACATTTCTCAAGCTGCTGATTTGCTCAAGTCCAGATCCAAATGTC 1099
Db 281 AsnLysItrThrThrPheLeuYrAlaArgLeuIleCysSertValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGATGTTTACGACTTCCAGTAACATTTTCAAGGATGACCGTG 1219
Db 321 AsnProValValYrGlyAlaPheThrThrSertAsnIlePheLysGlySertAlaVal 340
QY 1220 TGTATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1279
Db 341 CysMetIyrSertIserAspValAlaArgValPheLeuGlyProYrAlaHisArgAsp 360
QY 1280 GAGCCCAATATCAATGAGTGCCTTATCAAGAAAGATCCCATATCAGGCGCAGAGACT 1339
Db 361 GlyProAsnIyrdIntrpValProYrGlnGlyArgValProYrProArgProGlyYr 380
QY 1340 TGTCCAGAAACATTTGTGTGTTTGACTTCAAGAAAGACCTTCTGATGATGTTAT 1399
Db 381 CysProSertIserPheGlyGlyPheAspSertThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAGATATCCAGCATGATCATTCAGTGTCTTCTTATGAACATGCGCCA 1459
Db 401 ThrPheAlaArgSertIstProAlaMetYrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAAACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1519

Db 421 ILeValIleYsThInAspValAsnTrGInpHeThrGInIleValValAspArgValAsp 440
QY 1520 GCAGAAAGTGAACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGTTCTTAA 1579
Db 441 ALaGuaBpGlyGInTrAspValMetPheIleGlyThrAspValGlyThValLeuYs 460
QY 1580 GTAGTTCAATTCCTTAAGAGACTGTATGATTAGAAGAGGTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProYsGInThTrIpyrAspLeuGInGluValLeuLeuGInGluMet 480
QY 1640 ACAGTTTTCGGAAACGACTGCTATTTTCAACCAATGAGCTTTCACATAGCAGACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThlyGInGInGIn 500
QY 1700 CTATATATGTTGTAACAGCGCTGGGTTGCCAGCTCCCTTTAAACCGGTGATATTAC 1759
Db 501 LeuYrIleGlySerThrAlaGlyValAlaGInLeuProLeuHsArgCysAspIleYr 520
QY 1760 GGAAGAAGCGTGTGAGTGTGCTGCGCCGAGACCTTACTGTGCTTGGGATGTTCT 1819
Db 521 GlyYsAlaCysAlaGluCysCysLeuAlaArgAspProYrCysAlaIleTrAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAAGACGACAGACGACAAAGATATAAAT 1879
Db 541 AlaCysSerArgYrPheProThrAlaIleArgArgGlnArgGlnAspIleArgAsn 560
QY 1880 GGAAGACCACTGACTCACTGTTTCAAGACTTACACCAATGATATACCATGCGCAACGCTT 1939
Db 561 GlyAspProLeuTrHsAsCysSerAspLeuHsShsAspAsnHsShsGlyHsIleSerPro 580
QY 1940 GAAGAGAAATCATCTATGATGATAGAGAAATAGTACATATTTTGGAAATGAGTCCGAAG 1999
Db 581 GInGluArgIleIleYrGlyValGInAsnSerThrPheLeuGluCysSerProYs 600
QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGCGCAAAATGAAGCGCAAAAGAG 2059
Db 601 SerGlnArgAlaLeuValYrTrpGlnPheGlnArgAlaGlnGluValArgGluGln 620
QY 2060 ATCAGAGTGAATCATCTATCATCATCAGACAGACATCAAGGCTTCTGCTACGATGTTACA 2119
Db 621 IleArgValAlaAspAspHsIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGAGTTCAGGAATTAATCTGCGCATGCGAGGAGAAACATGGGTTCAATCAAACTCTT 2179
Db 641 GInYsAspSerGlyAsnTrLeuCYsHsIleAlaValGInHsGlyPheIleGlnThleu 660
QY 2180 CTTAAGGTAACTCGAAGTCAATGACACAGACATTTGAAGAATCTTTCATTAAGAT 2239
Db 661 LeuYsValThrLeuGluValIleAspThrGlnHsIleuGlnGluLeuHsIleYsAsp 680
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCAAATGACATGACCTTACGCAAGAG 2299
Db 681 AspAspGlyAspGlySerYsThlySerGluMetSerAsnSerMetThrProSerGlnYs 700
QY 2300 GTCTGATACAGAGACTTCATGACAGTCAATCAACCAACCCATCTCAACAGATGATGAG 2359
Db 701 ValTrpYrAspAspPheMetGlnLeuIleAsnHsIleProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAACAACTCGGCAAGGCGCAGACATACC 2419
Db 721 PheCYsGInGInValIleTrpYsArgAspArgYsGlnArgGlnArgGlnArgGlnArg 740
QY 2420 CCAGGAAACAGTAACTAATGAGAAGCTTACAAAGAAATTAAGAAAGTGAAGAAAGAGG 2479
Db 741 ProGlyAsnSerAsnLeuIleTrpYsHsIleuGlnGlnGlnGlnGlnGlnGlnGlnGln 760
QY 2480 ACCCAGCAATTTAGAGAGGACCCAGAGAGTGC 2512
Db 761 ThrHsGlnPheGlnArgAlaProArgSerVal 771

RESULT 4
US-08-833-391-54

Sequence 54, Application US/08833391
Patent No. 6013781
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-SEP-1993.
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-391-54
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-833-391-54 (1-771)
QY 200 ATGGGCGGTAACTAGATGTCGTGCTTTCGGGAGATATTACTTACGAGACGA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuTrpAlaArgAla 20
QY 260 AACTATCAGATGAGGAAGAACATATGTCACAGGCTGAATATTCTTACAAAGAAATGTTG 319
Db 21 AsnTrGlnAsnGlyYsAsnAsnValProArgLeuYleuSerYrIlyGInGluMetLeu 40
QY 320 GAATCAACAATGTGATCACTTCAATGCTTGCCCAACAGCTCCAGTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerYrHsIleThrPhe 60
QY 380 CTTTGGATAGGAAGGAGTAGGCTGTATGTTGAGCAAGGATCAATATTTCATTTC 439
Db 61 LeuLeuAspGluGlnArgSerArgLeuYrValGlyAlaYsAspHsIlePheSerPhe 80
QY 440 GACCTGGTAAATCAAGATTTTCAAAAGATTGTGCGCAGTATCTTACACCAAGAGA 499

Db 81 AspleVal)asnilleysaspheglnlyleValTrpValSerTyrThrArgArg 100
QY 500 GATGATATGCAAGTGGCTGGAAGAAAGACATCTGGAAGATGTGCTAAATTCATCAAGTA 559
Db 101 AspglucylsertrpalaiglylyasprileuleuylsglucysalaasnpheilleuVal 120
QY 560 CTTAAGCATTAATCAAGACTCACTTGTACCGCTGTGGAACGGGGCTTTTCATCAATT 619
Db 121 LeuylsalatYrAangIntnHileuYrAlaCysglYThrAlaPheHISProIle 140
QY 620 TGCACCTACATTTGAATTTGACATCATCTCGAGACAAATATTTTAAGCTGAGAACTCA 679
Db 141 CysThrYrIleGlnIleGlyHISISpRoglunaspasnillepelysleuGluAspSer 160
QY 680 CATTTGAAAACGGCGGTGGGAAGATCCATATGACCTTAAGCTGTGACAGCATCCCT 739
Db 161 HispheiunsnoglyargglylysserProlyrAspProlylsleuHnHISerleu 180
QY 740 TTAATAGATGAGAAATTAATCTCTGAACTGACGCTGATTTTATGGGGCAGACTTGTCT 799
Db 181 LeuileaspglyleuYrserglyThrAlaAlaaspPheMetGlyArgaspPheAla 200
QY 800 ATCTTCGGAATCTTGGGGACACCAACCAATCGAGACAGACAGCATGATTCAGGTGG 859
Db 201 IleasparThrleuGlyHISISpRoglunaspasnillepelysleuGluAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTCATTAAGTCCACCTCTCAGAGAGTGAACAATCCGTAAGAT 919
Db 221 LeuasnaspProlylspeHISerAlaHISleuHISerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTTCCGTGAATGCAATAGATGAGAAACATCTGGAAGAACT 979
Db 241 AspyValIYrPhePhePheArgGluAsnAlaIleaspglyHISISerGlylyAla 260
QY 980 ACTCAGCTGATGATAGTCAATGATGCAAGATGATCTTGGAGGGCACAAGTCTGGTG 1039
Db 261 ThrHISAlaArgIleGlyGlnIleCyslyAsnaspheglnlyHISArgSerleuVal 280
QY 1040 AATAAATGACAACATCTCTCAAGCTGCTGTGCTCAGTGCAGAGGTCACAAATGCG 1099
Db 281 AsnlyserTrpThrPheleuValaArgleuHISerSerValProGlyProAsnGly 300
QY 1100 ATTGACATCATTTTGAAGTGAAGTGAAGTGTATTCCTAATGAATTTAAAGATCCATAA 1159
Db 301 IleasprHISpHeaspGluLeuGlnAspValPheleuMetAsnPhelysaspProlys 320
QY 1160 AATCCAGTTTATATGAGTGTTCACATTCCTCAATCAATTTTCAAGGATCAGCGTG 1219
Db 321 AsnProValaIYrGlyValPheThrTrpSerSerAsnIlePheylsglySerAlaVal 340
QY 1220 TGTATGATAGCATGAGTGTGAGAAAGGTGTCTTGTGCATATGCGCCACAGGAT 1279
Db 341 CysMetCysSerSerSerAspValaArgValaPheleuGlyProlyrAlaHISArgAsp 360
QY 1280 GGAACCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTATCCACGGCCAGAGACT 1339
Db 361 GlyProAsnYrGlnTrpValProlyrGlnGlyArgValaProlyrProAsnProGlyYr 380
QY 1340 TGTCCCAAGCAAAATTTGGTGTCTTCACTTCAAAAGACCTTCTGTATGATTTATA 1399
Db 381 CysProserYrThrPheGlyGlyPheAspSerThrYsAspLeuProAspValIle 400
QY 1400 ACCTTTGACAGAGTCAATCAGCATGTATCAATCCAGTGTTCCTAATGAACATTCGCCA 1459
Db 401 ThrPheHISArgSerHISProAlaMetTrpAsnProAlaPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAAACGATGTAAATTAATCAATTTTACAAATTTGCTGAGCCGAGTGAT 1519
Db 421 IleValIleYrThrAspValaAsnYrGlnPheHISGlnIleValaValaAspArgValaAsp 440
QY 1520 GCGAAGATGAGACATGATGATTTATGTTATGCGAAGACATGTTGGAGCCGTTCTTAA 1579
Db 441 AlaGluAspIYrGlnYrAspValaMetPheIleGlyThrAspValaGlyThrValleuYs 460

QY 1580 GTAGTTTCATTCCTAAGAGACTTGATGATTTTAGAAGAGTTCTGCTGAAGAATG 1639
Db 461 ValValSerIleProlylsGlnTrpYrAspLeuGlnValleuLeuGlnGluMet 480
QY 1640 ACAAGTTTTCGGGAACCGACTGCTATTTTCAGAAATGAGCTTTCACATAAGACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrYsGlnGln 500
QY 1700 CTATATATGTTCAACCGCTGGGGTTCGCCACTCCCTTAAACCGGTGTATTTAC 1759
Db 501 LeuYrIleGlySerThrAlaGlyValaAlaGlnLeuProleuHISArgCysAspIleYr 520
QY 1760 GGGAAAGCGTCTGAGTGTCTGCTGCGCCGAGACCTTACTGTGTGGATGGTCT 1819
Db 521 GlyValaIYrAlaGluCysCysleuAlaArgAspProlyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGACAGACAGATTAAGAAT 1879
Db 541 AlaCysSerArgYrPheProThrAlaYsArgTrpArgArgGlnAspIleArgAsn 560
QY 1880 GGAAGCCCACTGACTACTGTTCAAGACTTAACCATGATATACCATGGCCACAGCCT 1939
Db 561 GlyAspProleuThrHISCysSerAspLeuHISISpAsnHISISGlyHISerPro 580
QY 1940 GAAGAGAGATCATCTATGTTGAGAGATAGAGACATTTTGAATGACATCCGAG 1999
Db 581 GluGluArgIleIleYrGlyValaGluAsnSerSerThrPheleuGluCysSerProlys 600
QY 2000 TCGCAGAGACCGTGTCTATTTGGCAATTCAGAGCGGAATGAGAGCGAAAGAAAG 2059
Db 601 SerGlnArgAlaLeuValYrTrpGlnPheGlnArgAsnGlnGluArgGlySgluGln 620
QY 2060 ATGAGAGTGAATATCATCATCAGACAGATCAAGCGCTTCTGCTAGTACTTACAA 2119
Db 621 IleArgValaAspAspHISISleIleArgTrpAspGlnGlyLeuLeuLeuArgSerleuGln 640
QY 2120 CAGAGAGATTCAGGCAATTAACCTGCTGCAAGCGGTGGAACATGGTTCATCAAACTCT 2179
Db 641 GlnlyAspSerGlyAsnYrleuGlyHISAlaValaGlnHISGlyPheIleGlnTrpLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCAATGACAGACATTTGGAAGACTTTCATTAAGAT 2239
Db 661 LeuYrValThrleuGlnValIleAspThrGlnHISleuGlnGluLeuLeuHISlyAsp 680
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATATGATGACACATTCACGAAG 2299
Db 681 AspAspGlyAspGlySerYsThrYsGluMetSerAsnSerMetThrProserGlnYs 700
QY 2300 GTCTGTACAGAGACTTCATGACGCTCAACACCAACCCCAATCTCAACAGATGAG 2359
Db 701 ValTrpYrArgAspPheMetGlnleuHISpRasnleuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGAAGAAACAAGTCTGGCAAGGCCAGCATACC 2419
Db 721 PheCysGlnGlnValTrpYsArgAspArglySgluArgGlnArgProGlyHISThr 740
QY 2420 CCAAGGACAGTACAAATGAGAAACACTTACAGAAATTAAGAAAGTGAAGACAGAG 2479
Db 741 ProGlyAsnSerAsnYrTrpYsHISleuGlnGlnAsnlyleGlyArgAsnArgArg 760
QY 2480 ACCCAAGATTTGAGAGGGGACCCAGAGAGTGC 2512
Db 761 ThrHISGluPheGlnArgAlaProArgSerVal 771

RESULT 5

US-09-060-610-54

; Sequence 54, Application US/09060610

; Patent No. 6344544

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey S.

; APPLICANT: Kolodkin, Alex L.

; APPLICANT: Matthes, David

APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-610-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-060-610-54 (1-771)

QY 200 ATGGCTGGTTAACTAGATGCTGCTTTCTGGGGAGTATTACTTAAGCAAGACA 259
DB 1 Mccgltprpneunrtrrvgllvalcyleubhetrrpdlvalleuleunthralarqala 20
QY 260 AACTATCAGATGGAGGAACAATGTGCGCAAGCTGAAATTATCTCAAGAAATGTG 319
DB 21 Asntrrglnsngllylsasnaasnvalprohrgleuylsleuserrlylsglmetleu 40
QY 320 GAATCCAAACATGTGATCACTTTCAATGGCTGGCCCAACAGCTCAAGTTATCACTTC 379
DB 41 GluserbsnbnvalilietnrphasnnglyleuallasnserSerrSYRhlsthphe 60
QY 380 CTTTGGATGAGGAACGAGTGGCTGTATGTGGAGCAAGATCAATTTTTCATTC 439
DB 61 Leuleunspglungrlungrsertrglnleuvalalvalalysapshsilepneserphe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAGAATGTGCGCCATGATCTTACCCAGAGA 499
DB 81 Aspleunvalasnlelysaapshpnehllylvalitrrprovalserlyrthrarg 100
QY 500 GATGATGCAAGTGGCTGGAAAAGACATCTGAAAAGATGTCTAATTTTCATCAAGCTA 559
DB 101 Aspgluncylsrlpalaiglylsasplleleuylsglucysalaasnphelilelyval 120

QY 560 CTTAAGCATATATACAGACTCACTTGACGCTGTGGAACGGGGCTTTTCACTCAATT 619
DB 121 LeuleyalaLyrrsnglnthrlsleuylalacylegylthrclalaphenlsprolle 140
QY 620 TGCACCTACATTTGAATTGGACATCATCTGAGACACATATTTTAACTGGAGAACTCA 679
DB 141 Cysrthrlyrllleglllegllylshlsrproglaspsanllepheylsleuglunaser 160
QY 680 CATTGGAACCGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGGTCAAGACATCCCT 729
DB 161 Hlsphneglunsngrlylrglylserserprolyraspprolyseuleunthralaserleu 180
QY 740 TTAATAGATGAGAATATATCTCTGGAACCTGACAGCTGATTTTATGGGGCCAGACTTGT 759
DB 181 Leulileaspcllygluleuylrsergylthralalalspshenctgylasapshenala 200
QY 800 ATCTTCGAACTCTTGGGACCAACCAACCAATGAGACAGACAGCATGATTCAGAGTGG 859
DB 201 Hlsphetrgrthrleuglylshlsrprolleargthrclunghlsaspserrargtrp 220
QY 860 CTCGAATGATCAAGTTCAATTAGTGGCCACCTCATCTCAGAGAGTGACATCTGGAAGAT 919
DB 221 Leunsnspprolysphelelserralshlsleuileserglunseraspshnproglunsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACACTCTGAAAAAGCT 979
DB 241 Aspilyvallyrrhephepnearglunaslalaleaspcllylunhsersgyllylala 260
QY 980 ACTCAGCTTGAATGAGTCAATATGCAAGATGACTTTGGAGGGCAGAAAGCTGTGTG 1039
DB 261 Thrhisalarglilegylnglilecyllysasnaspshedlygllylshserserleuval 280
QY 1040 AATAATGAGCAACATTCCTCAAGCTCGCTGATTTGCTCAGTGCAGAGTCCAAATGGC 1099
DB 281 Asnlystrpthrthrphneulysalargyleuilecyservalproglproasnrgly 300
QY 1100 ATTGACATCATTTTGTATGATGACAGTGAAGTATTCCTAATGAACTTTAAAGATCTTAA 1159
DB 301 Hlsapthrhlshpshaspglunleuglnasppvalpheleunserbnheylsaspproly 320
QY 1160 AATCCAGTTGTATATGAGAGTGTTCACGACTTCCAGTAACTTTTCAAGGATCAAGCTGT 1219
DB 321 Asnprovalalylrglylvalpethrthrnserserasnlephelysglserralaval 340
QY 1220 TGTATGATATGATGATGATGATGAGAGAGTGTCTGTCATATGAGCCACAGAGAT 1279
DB 341 Cysmectylsermetseraspvalargargvalpheleuglylprotyrhlshlsargasp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTATCCAGCGCCAGAACT 1339
DB 361 Glyproasnlyrglntrpvalprotyrglnglylvalprotyrprohargprogllythr 380
QY 1340 TGTCCAGCAAAAATTTGGTGTGTGACTTACCAAGAGACCTTCTGATGATTTAT 1399
DB 381 CysproserlystrhrhegllylphespserrThrlyasppneuproaspasvalille 400
QY 1400 ACCTTGGCAAGATGATCAGGCATGTAACAATCCAGTGTTCCTATGAAATTCGCCCA 1459
DB 401 Thrphelalargserrhsproalawetlyraspprovalpheprohewcshnanaargpro 420
QY 1460 ATAGTATCAAAAACGATGTAAATTTATCAATTTTACAAATTCGTAGACCGAGTGAT 1519
DB 421 HlevalillelystrhrspvalasnlyrglnpheThrGlnlevalalaspargvalasp 440
QY 1520 GCAGAAAGATGACAGTATGATGTTATGTTATGGAACAGATTTGGAGCCGTTCTTAA 1579
DB 441 Alaglunspcllyluntyraspvalmetphelleglythrasyalaglythrvalleuys 460
QY 1580 GTAGTTTCATTTCTTAAAGAGACTTGTGATGATTTTGAAGAGGTTCTGCTGGAAGAATG 1639
DB 461 Valvalserlleprolysglunthrrptryraspleuglunghlvalleuueuglunghmet 480
QY 1640 ACAGTTTTTCGGGAACCGACTGCTATTTCAGCAATGAGAGCTTTCACATAAGCAGAAACA 1699

```

Db      .481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrIlyGlnGln 500
QY      1700 CTATATATTGGTTCAACGGCTGGGGTTGGCCAGCTCCCTTTACACGGGTGTGATATTAC 1759
Db      501 LeuTyrlIleGlySerThrAlaGlyAlaIleGlnLeuProLeuNH2ArgCysAspIleTy 520
QY      1760 GGGAAAGGCGTGGCTGAGTGTGTTCCTCCGCCGAGACCTTTACTGTGCTTGGGATGGTTCT 1819
Db      521 GlyIlybAlaCysAlaGluCysGlyLeuAlaArgProTyCysAlaTrpAspGlySer 540
QY      1820 GCATTTCTCCGCTATTTTCCCACTCCAAAGAGACCAAGACCGCAAGATATATAGAAAT 1879
Db      541 AlaCysSerAlaGlyTyPheProThrAlaLysArgGlnArgGlnAspIleArgAsn 560
QY      1880 GGAGACCCAGTCACTCACTGTGTTCAGACTTACACATGATTAATCAACATGAGCCACAGCCCT 1939
Db      561 GlyAspProLeuThrIleCysSerAspLeuNH2AspAsnNH2IleGlyNH2SerPro 580
QY      1940 GAAGAGAGATCATCTTATGTGTGAGATAGATATGTAGACATTTTGGATTCAGATCCGAG 1999
Db      581 GlnGluArgIleIleTyGlyValGlnLysSerSerThrPheLeuGluCysSerProLys 600
QY      2000 TCGCAGAGAGGCGTGTATTGGCAATTCAGAGGCGCAATGAAAGCGCAAGAAAGAGAG 2059
Db      601 SerGlnArgAlaLeuValTyTrpGlnPheGlnArgGlnArgGlnArgGlnGln 620
QY      2060 ATCAGAGTGAATGATCATATATCATCAGACACATCAAGGCGCTTGTGCTACTAGTCTCAA 2119
Db      621 IleArgValaAspAspNH2IleIleArgThrArgGlnGlyLeuLeuLysSerLeuGln 640
QY      2120 CAGAAAGATTCAAGCAATTACCTCTGCCATGCGGTGGAACATGGGTTCATACAACCTCT 2179
Db      641 GlnLysAspSerGlyAsnTyLeuCysNH2IleAlaValGlnIleGlyPheIleGlnThrLeu 660
QY      2180 CTTAAGGTAAACCTCGAAGTCATTTGACACAGAGCATTTGGAAAGAACTTTCATTAAGAT 2239
Db      661 LeuLysValaIleIleGlnValaIleAspThrGlnIleGlnGluLeuLeuNH2IleAsp 680
QY      2240 GATGATGAGATGCGCTTAAAGCCAAAGAAATGTCCAATATAGACATGACACCTAGCAGAG 2299
Db      681 AspAspGlyAspArgIleSerIleThrIlybGlnMetSerAsnSerMetIleProSerGlnLys 700
QY      2300 GTCTGTGTCAGAGACTTCATGATCAGCTCATCAACCAACCCCAATCTCAACACGATGATGAG 2359
Db      701 ValTrpTyArgAspPheMetGlnLeuNH2LeuNH2ProAsnLeuAsnThrMetAspGlu 720
QY      2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAACAAAGCTGGCGCAAGGCCAGACATACC 2419
Db      721 PheCysGluGlnValaIleTrpLysArgAspArgLysGlnArgGlnArgProGlyNH2Thr 740
QY      2420 CCAGGAAACAGTAAACAATGGAGAGACTTACAAGAAATTAAGAAAGGTAGAAACAGAGAG 2479
Db      741 ProGlyAsnSerAsnLysTrpLysNH2IleuGlnGlnLysLysLysGlyArgAsnArgArg 760
QY      2480 ACCCAGCAATTGAGAGGGCACCAGAGAGTGC 2512
Db      761 ThrIleGluPheGlnArgAlaProArgSerVal 771

RESULT 6
PCT-US94-10151A-54
; Sequence 54, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/10151A
: FILING DATE: 13-SEP-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: Pp-58750-PC/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELE: 910 277299 FHT UR
: INFORMATION FOR SEQ ID NO: 54:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 771 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US94-10151A-54

Alignment Scores:
Pred. NO.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US94-10151A-54 (1-771)

QY 200 ATGGCGTGTAACTGACATGCTGTCTGTCCTTTCTGGGAGATTAATTACACGACAAAGCA 255
Db 1 MetGlyTyrPheuhirxrgilevalCysleuphetrgilyValleuLeuthralaargAla 20

QY 260 AACATACAGAAATGGGAAAGAAACAATGGCCCAAGGCGTGAATAATTCCTCAAGAAATGTTG 319
Db 21 AsntryGlnhsmcilyushsmasnvalProhagleuylsleuSerTyrLysGluMetLeu 40

QY 320 GAATCCAAACATGTGATCACTTTCATATGGCTTGCCCAACGCTCAGTTATCATACCTTC 379
Db 41 GluSerhshsnvalIlethrPhehsmgilyleuAlaenserserTyrhIsthrPhe 60

QY 380 CTTTGGATGAGGACGAGAGTNGGCTGATGTTGGAGCAAAAGATCACATATTTTCATTG 439
Db 61 LeuLeuAspGluGluargSerargLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTAAATATCAAGATTTTCAAAAGATGTGTGGCCAGATATCTTACACAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnIlyIleValIleProValSerTyrThrArgArg 1000

QY 500 GARGAATGCAAGTGGGCTGAAAAGACATCCTGAAAGAATGTGTAATTTATCATCAAGTA 559
Db 101 AspGluCylbelystrPalaGilyLysAspIleleuLysGluCysAlaIenPheIleVal 1200

QY 560 CTTAAGGCAATATATCAGACTCACTTGTAACGCTGTGGAAACGGGGCTTTTCATCAATT 619
Db 121 LeuLysAlaIyLAsnGlnIthrHisIleuTyrAlaCysGlyIthrGlyAlaPheHisProIle 1400

QY 620 TGACCTACATTTGAATTTGGAGATCATCTCGAGACAAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisIsthrProGluAspAsnIlePheLysLeuGluAsnSer 1600

QY 680 CATTTTGAAAACGGCGGTGGAAAGAGTCATATGACCTTAAGCTGCTGACAGATCCCTT 739
Db 161 HisPheGluAsnGilyArgGilyLysSerProTyrAspProLysLeuLeuIthrAlaSerLeu 1800

QY 740 TTAATGATGAGAAATTAATTACTCTGAAATCGACGTGATTTTAATGGGCGAGACTTTGCT 799

```


DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-3 (1-655)

QY 200 ATGGCGTGTAACTAGAGTGTCTGCTTTCTGGGAGATTAATCTTACAGACAGACA 259

DB 1 MetcyltrpLeuThrAlaValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTACAGAAATGGAGAAACAATGTCCAGAGCTGAAATTAATCTTCAAGAAATGTG 319

DB 21 AsnIyrcIlnAsnGlyIysAsnAsnValProArgLeuIysLeuSerIyrlYsGluMetLeu 40

QY 320 GAATCCAAACAATGTGATCACTTTCATAGCTGTGGCCCAACGCTCCAGTATCATACCTTC 379

DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerIyrlYsIleThrPhe 60

QY 380 CTTTGTGATGAGAGAGAGAGCTGTATGCTTGAGCAAAAGATTCATATTTTTCATTC 439

DB 61 LeuLeuAspGluGluAspArgLeuIyrlValGlyAlaIysAspHisIlePheSerPhe 80

QY 440 GACCTGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGA 499

DB 81 AspLeuValAsnIleIysAspPheGlnIysIleValIleTrpValSerIyrlYsIleThrArgArg 100

QY 500 GATGATGCAAGTGGCTGAGAAAGACATCTCTGAAGAAATGTGCTAAATTTTCATCAAGTA 559

DB 101 AspGluCysIyrlYsIleTrpAlaGlyIysAspIleLeuIysGluCysAlaAsnPheIleIysVal 120

QY 560 CTTAAGGCATATATCATGACTCTTGTATGCTGTGGAACGGGGGCTTTTCATCCAT 619

DB 121 LeuIysAlaIyrlYsAsnGlnThrHisIleuIyrlAlaCysGlyIleThrGlyAlaPheHisIleProIle 140

QY 620 TGCACCTACATTTGAAATTTGACATCATCTGAGAGACAATATTTTAACTGGAGAACTCA 679

DB 141 CysThrIyrlIleGluIleGlyIleHisIleAspGluAspAsnIlePheIysLeuGluAsnSer 160

QY 680 CATTTTGAACCGCCGTGGAGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCT 739

DB 161 HisPheGluAsnGlyIyrlYsSerProIyrlAspProIyrlAsnLeuThrAlaSerLeu 180

QY 740 TTAATGATGAGAAATTAATCTCTGGAATCTGCACTGATTTTATGGGGCCAGACTTGTCT 799

DB 181 LeuIleAspGlyIyrlYsLeuIyrlSerGlyIleThrAlaIleAspPheMetGlyIyrlAspPheAla 200

QY 800 ATCTCCGAATCTGTGGGACACCAACCAATCAGACAGACAGACAGATGATTCAGAGTGG 859

DB 201 IlePheArgThrIleuGlyIleHisIleSerProIleArgThrIleGlnHisIleAspSerArgTrp 220

QY 860 CTCATGATGCAAAAGTTCATTAGTCCCACTCATCTCAGAGAGTGAACAATCTCTGAAGAT 919

DB 221 LeuAsnAspProIyrlYsPheIleSerAlaHisIleuIleSerGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTATATCTTTTCTTCGCGTGAATGCAATGATGAGAAACACTCTGGAAGACT 979

DB 241 AspIysValIyrlYsPhePheArgGluAsnAlaIleAspGlyIyrlYsIleSerGlyIysAla 260

QY 980 ACTCAGCCTGAGATAGATGATGCAAGATGACCTTGGAGGGCAACAAGCTGGCTG 1039

DB 261 ThrHisIleAlaArgIleGlyIleIleCysIysAsnAspPheGlyIyrlYsIleSerLeuVal 280

QY 1040 AATTAATGACAACAATCTCTCAAGACTGTCTGATTTGTCTCAGTGCAGGCTCCAAATGGC 1099

DB 281 AsnIyrlYsIleThrPheLeuIyrlYsAlaArgLeuIleIysSerValProGlyIleProAsnGly 300

QY 1100 ATTGACACTCATTTTGTGATGACTGCGAGATGTATCTTAATGAACCTTAAAGTCTTAA 1159

DB 301 IleAspThrHisIlePheAspGluLeuGlnAspValPheLeuMetAsnPheIysAspProIyrlYs 320

QY 1160 AATCAGATTGATATGAGATGTTAAGACTTCCAGTAACTTTTCAAGGGATGAGCGCG 1219

DB 321 AsnProValValIyrlYsIleValIlePheThrHisSerSerAsnIlePheIysGlyIleSerAlaVal 340

QY 1220 TGTATGATATGATGATGATGATGAGAGGGGTGTCTTGTGTCATATGCCCACAGGGAT 1279

DB 341 CysMetIyrlSerMetSerAspValAlaArgValPheLeuGlyIyrlYsIleThrAlaHisIleArgAsp 360

QY 1280 GGACCCCAATCATTAATGGGGGCTTATCAAGAAAGCTCCCTTATCAACGCCCAAGCAACT 1339

DB 361 GlyProAsnIyrlYsIleThrValProIyrlYsGlnIyrlYsAlaProIyrlYsProAspProGlyIyrlYs 380

QY 1340 TGTCCGACCAAAACATTTGGTGTGCTTGAATCTCAACAAAGACCTTCTGATGATGATATA 1399

DB 381 CysProSerIyrlYsIleThrPheGlyIyrlYsPheAspSerThrIysAspLeuProAspAspValIle 400

QY 1400 ACCTTTGCAAGAAATGATATCCAGCCATGATACATTCAGTGTTCCTATGAACATCGCCA 1459

DB 401 ThrPheAlaArgSerHisIleProAlaMetIyrlYsAspProValPheProMetAsnAsnArgPro 420

QY 1460 ATAGTATCAAAACGATATGATTAATTCATTTTCAACAAATTTGCTGATGACCGAGTGGAT 1519

DB 421 IleValIleIyrlYsThrAspValAsnIyrlYsGlnPheThrGlnIleValIleAspArgValAsp 440

QY 1520 GCAGAAAGATGACAGATATGATGTTATGTTATGGAACAGATGTTGGACCGTCTTAA 1579

DB 441 AlaGluAspGlyIyrlYsIleThrAspValMetPheIleGlyIleThrAspValGlyIleThrValIleuIys 460

QY 1580 GATGTTCAATTTCTTAAGAGACTTGCTATGATTTGAAGAGGTTGTCTGTAAGAAATG 1639

DB 461 ValIleSerIleProIyrlYsGluThrTrpIyrlYsAspLeuGluValLeuLeuGluIyrlYsMet 480

QY 1640 ACAGTTTTTGGGAAACCGACTGATATTTCAAGCAATGAGCTTTCCACTAAGCAGCAACA 1699

DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrIyrlYsGlnGln 500

QY 1700 CTATATATGTTGTTCAACGGCTGGGTTGCGCCAGCTCCCTTACACGGGTGTGATTTTAC 1759

DB 501 LeuIyrlIleIyrlYsThrIleGlyIyrlYsAlaGlnIleProIyrlYsIleArgCysAspIleIyrlYs 520

QY 1760 GGGAAACGCTGTGCTGATGTTGCTCGCCCGAGACCTTACTGTGCTGGAGTGGTCT 1819

DB 521 GlyIyrlAlaCysAlaGlyIyrlYsCysIleuAlaArgAspProIyrlYsAlaIleTrpAspGlyIyrlYs 540

QY 1820 GCATTTTCTGCTATTTTCCACTGCAAGAGACGCAACAGCAACAAGATTAAGAAAT 1879

DB 541 AlaCysSerArgIyrlYsPheProThrAlaIyrlYsArgIyrlYsGlnAspIleArgAsn 560

QY 1880 GGAGACCACTGATCTGATGTTTCACTTACACCATGATATACCATGAGCCCAACGCT 1939

DB 561 GlyAspProLeuThrHisIyrlYsSerAspLeuHisIleAspAsnHisIleGlyIleSerPro 580

QY 1940 GAAAGAGATCATCTATGCTGTAGAGAAATAGTACCAATTTTGGATGCACTGCGAAG 1999

DB 581 GluGluArgIleIleIyrlYsAlaGluAsnSerThrPheLeuIyrlYsSerProIyrlYs 600

QY 2000 TGCAGAGACCGCTGTCTATTTGGCAATTCAGAGCGCAATGAGCGGAAAGAGAG 2059

DB 601 SerGlnArgAlaLeuValIyrlYsTrpGlnPheGlnArgIyrlYsGlnIyrlYsGluGln 620

QY 2060 ATCAGAGTGAATGATATCATTCAGAGACAGATCAAGGCTTCTGATGATGATCAACA 2119

DB 621 IleArgValAspAspHisIleIleArgIyrlYsIleArgIyrlYsLeuLeuLeuIyrlYsSerLeuGln 640

QY 2120 CAGAAAGATTCAGGCAATTAATCTGCGATGCGGTGAGCAATGAG 2164

DB 641 GlnIyrlAspSerGlyIyrlYsAsnIyrlYsIleAlaValGluHisGly 655

RESULT 8

US-08-136-922-2

; Sequence 2, Application US/08136922

; Patent No. 5416197

; GENERAL INFORMATION:

; APPLICANT: Rapar, Jonathan A.

; APPLICANT: Luo, Yuling

; TITLE OF INVENTION: Compositions Which Regulate Neural

; TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same

; NUMBER OF SEQUENCES: 2

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 775
TYPE: PRT
ORGANISM: Mus sp:
US-09-308-1798-1
Alignment Scores:
Pred. No.: 0.000582 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-308-1798-1 (1-775)
QY 1778 TGTTCCTGCGCCGAGACCTTACTGCTTGCGATGCT 1816
DB 529 CysCysLeuAlaArgAspProtyrCyAlaTrpaspGly 541
RESULT 10
US-09-041-236-4
Sequence 4, Application US/09041236
Patent No. 6225285
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
TITLE OF INVENTION: Semaphorin K1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,236
FILING DATE: March 11, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-236-4
Alignment Scores:
Pred. No.: 0.466 Length: 606
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-041-236-4 (1-606)
QY 1043 AAATGACAACATTCTCAAGCTGCTG 1072
DB 1043 AAATGACAACATTCTCAAGCTGCTG 1072

DB 255 LysTrpThrThrPheLeuLysAlaArgLeu 264
RESULT 11
US-09-771-467C-4
Sequence 4, Application US/09771467C
Patent No. 6583277
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1 Polypeptides
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,467C
FILING DATE: 26-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-771-467C-4
Alignment Scores:
Pred. No.: 0.466 Length: 606
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-771-467C-4 (1-606)
QY 1043 AAATGACAACATTCTCAAGCTGCTG 1072
DB 255 LysTrpThrThrPheLeuLysAlaArgLeu 264
RESULT 12
US-09-341-461-34
Sequence 34, Application US/09341461
Patent No. 6586389
GENERAL INFORMATION:
APPLICANT: Hammond, Timothy G.
Verruust, Pierre J.
TITLE OF INVENTION: Cubilin Protein, DNA Sequences Encoding Cubilin
FILE REFERENCE: D6148
CURRENT APPLICATION NUMBER: US/09/341,461
CURRENT FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: PCT/US99/01259
PRIOR FILING DATE: 1999-01-21
NUMBER OF SEQ ID NOS: 40
SEQ ID NO 34

```
; LENGTH: 112
; TYPE: PRT
; ORGANISM: human
; FEATURE:
; OTHER INFORMATION: amino acid sequence of C1s CUB1 domain
US-09-341-461-34

Alignment Scores:
Pred. No.: 5.27      Length: 112
Score: 9.00          Matches: 9
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.03%      Indels: 0
DB: 4                Gaps: 0

US-09-774-490-1 (1-2709) x US-09-341-461-34 (1-112)
QY 151 TTGTGCGCCGAGAGAGTTCACACAT 125
Db 67 LeuCYcgIyGlnArgSerSerAsnAsn 75

RESULT 13
US-08-556-422A-6
; Sequence 6, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Gallus gallus
US-08-556-422A-6

Alignment Scores:
Pred. No.: 4.69      Length: 295
Score: 9.00          Matches: 9
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 4                Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-6 (1-295)
QY 1784 CTGCGCCGAGAGACCTTACTGTGCTTG 1810
Db 287 LeuAlaArgAspProTyrCysAlaItrp 295

RESULT 14
US-08-556-422A-7
; Sequence 7, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 7
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-556-422A-7

Alignment Scores:
Pred. No.: 4.48      Length: 425
Score: 9.00          Matches: 9
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 4                Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-7 (1-425)
QY 1055 TTCTCAAGCTGTCTGATTTGCTCA 1081
Db 104 PheLeuYsAlaArgLeuIleCysSer 112

RESULT 15
US-08-556-422A-4
; Sequence 4, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-556-422A-4

Alignment Scores:
Pred. No.: 4.29      Length: 607
Score: 9.00          Matches: 9
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 4                Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-4 (1-607)
QY 1784 CTGCGCCGAGAGACCTTACTGTGCTTG 1810
Db 483 LeuAlaArgAspProTyrCysAlaItrp 491

RESULT 16
US-09-252-991A-32350
; Sequence 32350, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32350
```



```
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-152-060-58

Alignment Scores:
Pred. No.: 52.7      Length: 57
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-152-060-58 (1-57)

QY 716 CTTAGCTGCTGACAGCATCCCTT 739
DB 41 Prolysleuthralaserleu 48

RESULT 20
US-09-252-991A-24681
; Sequence 24681, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24681
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24681

Alignment Scores:
Pred. No.: 47.2      Length: 139
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.92%      Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-24681 (1-139)

QY 1539 TCATCTGTCATCTTCTGCATCC 1516
DB 105 SerTyrcyproserSerAlaser 112

RESULT 21
US-09-252-991A-32158
; Sequence 32158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32158
; LENGTH: 132
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32158

Alignment Scores:
Pred. No.: 42.6      Length: 322
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-32158 (1-322)

QY 1315 AGTCCCTATTCAGCGCCAGGAGAC 1338
DB 242 SerProleuserThralaagaen 249

RESULT 22
US-09-252-991A-20193
; Sequence 20193, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20193
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20193

Alignment Scores:
Pred. No.: 42.1      Length: 360
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-20193 (1-360)

QY 1504 CGTAGCCGAGTGATGACAGAGA 1527
DB 22 ArgArgProserdrgCysArgArg 29

RESULT 23
US-08-464-148-2
; Sequence 2, Application US/08464148
; Patent No. 5710026
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,148
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/385,500
FILING DATE: 08-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-21
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-464-148-2

Alignment Scores:
Pred. No.: 41.9 Length: 374
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-464-148-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613
DB 282 Pheleuargargleuglymetile 289

RESULT 24
US-08-385-500-2
Sequence 2, Application US/08385500
Patent No. 571217
GENERAL INFORMATION:
APPLICANT: Sprecher, Cindy A.
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,500
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-385-500-2

Alignment Scores:
Pred. No.: 41.9 Length: 374
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-385-500-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613
DB 282 Pheleuargargleuglymetile 289

RESULT 25
US-08-846-784-2
Sequence 2, Application US/08846784
Patent No. 5747645
GENERAL INFORMATION:
APPLICANT: Sprecher, Cindy A.
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,784
FILING DATE: 30-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/385,500
FILING DATE: 08-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-846-784-2

Alignment Scores:
Pred. No.: 41.9 Length: 374
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-846-784-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613

Db 282 Pheleuargyleuglymeille 289
RESULT 26
US-09-080-044-4
; Sequence 4, Application US/09080044
; Patent No. 6074649
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe F.
; APPLICANT: BAUDU, Philippe G.
; APPLICANT: RIVIERE, Michel A.
; TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
; TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
; FILE REFERENCE: PERITONITIS
; CURRENT APPLICATION NUMBER: US/09/080,044
; EARLIER FILING DATE: 1998-05-15
; EARLIER APPLICATION NUMBER: PCT/FR96/01830
; EARLIER FILING DATE: 1996-11-19
; EARLIER APPLICATION NUMBER: 95/14450
; EARLIER FILING DATE: 1995-11-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-09-080-044-4
Alignment Scores:
Pred. No.: 40.9 Length: 454
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-080-044-4 (1-454)
Cy 2498 CCCCTCAATTGCGGTCTCC 2475
Db 317 Proserginilearglyser 324
RESULT 27
US-09-531-857A-4
; Sequence 4, Application US/09531857A
; Patent No. 6387376
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe F.
; APPLICANT: BAUDU, Philippe G.
; APPLICANT: RIVIERE, Michel A.
; TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
; TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
; FILE REFERENCE: PERITONITIS
; CURRENT APPLICATION NUMBER: US/09/531,857A
; CURRENT FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 09/080,044
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: PCT/FR96/01830
; PRIOR FILING DATE: 1996-11-19
; PRIOR APPLICATION NUMBER: 95/14450
; PRIOR FILING DATE: 1995-11-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-09-531-857A-4
Alignment Scores:
Pred. No.: 40.9 Length: 454

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-531-857A-4 (1-454)
Cy 2498 CCCCTCAATTGCGGTCTCC 2475
Db 317 Proserginilearglyser 324
RESULT 28
US-09-252-991A-25962
; Sequence 25962, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25962
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25962
Alignment Scores:
Pred. No.: 39 Length: 673
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-252-991A-25962 (1-673)
Cy 1863 CGCTTGCGCTCTTGCAGTG 1840
Db 443 Argdeuvalargleuphealaval 450
RESULT 29
US-09-157-257-4
; Sequence 4, Application US/09157257
; Patent No. 6375954
; GENERAL INFORMATION:
; APPLICANT: DUTTA, Sukanta K.
; APPLICANT: BISWAS, Biswajit
; APPLICANT: VENUGOPAL, Ramesh
; TITLE OF INVENTION: A SIZE-VARIABLE STRAIN-SPECIFIC PROTECTIVE ANTIGEN FOR
; TITLE OF INVENTION: POTOMAC HORSE FEVER
; FILE REFERENCE: 8172-9016
; CURRENT APPLICATION NUMBER: US/09/157,257
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/059,252
; EARLIER FILING DATE: 1997-09-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Ehrlichia risticii
US-09-157-257-4
Alignment Scores:
Pred. No.: 37.9 Length: 849

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-157-257-4 (1-849)

QY 515 GCTGAGAAAGACATCTGAAAGAA 538
Db 63 AAGlylysaapleleuylsglu 70

RESULT 30
US-08-644-456-1
Sequence 1, Application US/08644456
Patent No. 5747650
GENERAL INFORMATION:
APPLICANT: Kulesz-Martin, Molly F.
TITLE OF INVENTION: p53ae PROTEIN AND ANTIBODY THEREFOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dunn & Associates
STREET: P.O. Box 96
CITY: Newfane
STATE: New York
COUNTRY: U.S.A.
ZIP: 14108
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage
COMPUTER: Victor 300 SX/25 (IBM PC Compatible)
OPERATING SYSTEM: MS-DOS Version 5.0
SOFTWARE: Wordstar Professional Release 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/644,456
FILING DATE: 10-May-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,496
FILING DATE: 2-Aug-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Dunn, Michael L.
REGISTRATION NUMBER: 25,330
REFERENCE/DOCKET NUMBER: RPP:135C US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 433-1661
TELEFAX: (716) 433-1665
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17
TYPE: amino acids
STRANDEDNESS: n/a
TOPOLOGY: n/a
MOLECULE TYPE: peptide
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: n/a
ORIGINAL SOURCE:
ORGANISM: mouse
STRAIN: n/a
INDIVIDUAL ISOLATE: n/a
DEVELOPMENTAL STAGE: n/a
HAPLOTYPE: n/a
TISSUE TYPE: n/a
CELL TYPE: n/a
CELL LINE: n/a
ORGANELLE: n/a
IMMEDIATE SOURCE: sequenced from cDNA clone from mouse
LIBRARY: plasmid p6.3
IMMEDIATE SOURCE: epidermal cell RNA, Genbank Accession #M13874
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 11

MAP POSITION: p53 gene
UNITS:
FEATURE: n/a
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Kulesz-Martin et al.
TITLE: Endogenous p53 Protein Generated from Wild
TITLE: Type Alternatively Spliced P53 RNA in Mouse
JOURNAL: Mol. Cell. Biol.
VOLUME: 14
ISSUE: 3
PAGES: 1698-1708
DATE: March, 1994
AUTHORS: Han, K.A. and Kulesz-Martin, M.F.
TITLE: Alternatively Spliced p53 RNA in Transformed
TITLE: and No. 5747650mal Cells of Different Tissue Types
JOURNAL: Nucleic Acids Res.
VOLUME: 20
ISSUE: 8
PAGES: 1979-1981
DATE: 1992
AUTHORS: Arai, N. et al.
TITLE: Immunologically distinct p53 molecules generated
TITLE: by alternative splicing
JOURNAL: Mol. and Cell. Biol.
VOLUME: 6
ISSUE:
PAGES: 3232-3239
DATE: 1986
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-644-456-1

Alignment Scores:
Pred. No.: 562 Length: 17
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-644-456-1 (1-17)

QY 1302 CTTATCAGAGAGAGTCCCT 1322
Db 9 leuilelysgluuuserfro 15

RESULT 31
US-08-164-151-22
Sequence 22, Application US/08164151
Patent No. 5519115
GENERAL INFORMATION:
APPLICANT: Mepelli Ph.D. Claudio
APPLICANT: Derobertis, Cathy
APPLICANT: Stahl, Jerry
APPLICANT: Swerdloff Ph.D. Michael D.
APPLICANT: William Ph.D. Jon I.
APPLICANT: Everett Ph.D. Nicholas P.
APPLICANT: Bascomb Ph.D. Newell
TITLE OF INVENTION: Reverse Antimicrobial Peptides,
TITLE OF INVENTION: Antimicrobial Oligopeptides and Other Antimicrobial
TITLE OF INVENTION: Compositions, and Methods of Their Production and Use.
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
STREET: 600 South Ave., West, Suite 300
CITY: Westfield

STATE: New Jersey
COUNTRY: United States of America
ZIP: 07090
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/164,151
FILING DATE: 09-DEC-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,784
FILING DATE: 01-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teechner Esq., Michael H
REGISTRATION NUMBER: 32,862
REFERENCE/DOCKET NUMBER: Enlmont 3.0-042
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-654-5000
TELEFAX: 908-654-7866
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-164-151-22

Alignment Scores:
Pred. No.: 529 Length: 28
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-164-151-22 (1-28)

OY 182 TCCTTGTCTGCTTACTGCTT 162
Db 6 SerleucysCysphaserleu 12

RESULT 32
US-09-574-377-27
Sequence 27, Application US/09574377
Patent No. 6573370
GENERAL INFORMATION:
APPLICANT: La Du, Bert
TITLE OF INVENTION: PONS And Uses Thereof
FILE REFERENCE: UM-04408
CURRENT APPLICATION NUMBER: US/09/574,377
CURRENT FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 37
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-574-377-27

Alignment Scores:
Pred. No.: 511 Length: 37
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-574-377-27 (1-37)

OY 511 CTGCATTCATCTCTTCTGCT 491
Db 12 LeuAlaPheIleSerGly 18

RESULT 33
US-09-205-258-296
Sequence 296, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06

```

; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 296
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (43)
; OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-296

Alignment Scores:
Pred. No.: 502 Length: 43
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-205-258-296 (1-43)

QY 2253 CCATCTCATCATCATCTTTA 2233
Db 18 ProSerProSerSerSerSer 24

RESULT 34
US-09-369-247-82
; Sequence 82, Application US/09369247
; Patent No. 6563992
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 44 Human Secreted Proteins
; FILE REFERENCE: P2024P1
; CURRENT APPLICATION NUMBER: US/09/369,247
; CURRENT FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,157
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,137
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,341
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,141
; EARLIER FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals stop translation
US-09-369-247-82

Alignment Scores:
Pred. No.: 497 Length: 46
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-369-247-82 (1-46)

QY 583 GTGAGTCTGATTATATGCTT 563
Db 3 ValSerLeuIleLeuCysLeu 9

RESULT 35
US-09-489-847-175
; Sequence 175, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-175

Alignment Scores:
Pred. No.: 495 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-489-847-175 (1-48)

QY 2634 GAATTCAGCTGAGTTCACCC 2654
Db 30 GuileGlnLeuSerSerThr 36

```

RESULT 36
US-09-328-352-7261
; Sequence 7261, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7261
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7261

Alignment Scores:
Pred. No.: 480 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-328-352-7261 (1-62)

QY 514 CCACTGATTCATCTCTTCT 494
Db 37 ProleuAlaphelleSerSer 43

RESULT 37
5320958-24
; Patent No. 5320958
; APPLICANT: INOUYE, SUNIKO; HSU, MEI-YIN; EAGLE, SUSAN;
; INOUYE, MASAYORI
; TITLE OF INVENTION: ISOLATED BACTERIAL REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/315,316
; FILING DATE: 24-FEB-1989
; SEQ ID NO: 24
; LENGTH: 62
5320958-24

Alignment Scores:
Pred. No.: 480 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 6

US-09-774-490-1 (1-2709) x 5320958-24 (1-62)

QY 2185 CTTAAGAAGTGTGTATGAA 2165
Db 3 LeuylslySerLeuylrGln 9

RESULT 38
US-09-107-532A-4497
; Sequence 4497, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham

STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Denek
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4497:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...67
SEQUENCE DESCRIPTION: SEQ ID NO: 4497:
US-09-107-532A-4497

Alignment Scores:
Pred. No.: 475 Length: 67
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-107-532A-4497 (1-67)

QY 2462 TCTTATTTCTGTAGTGCT 2442
Db 15 SerTyrrheleuylSerAla 21

RESULT 39
US-08-456-647B-31
; Sequence 31, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-31

Alignment Scores:
Pred. No.: 474 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-456-647B-31 (1-68)

QY 1175 GGAGTGTTCAGACTTCAGT 1195
DB 55 GlyValPheThrSerSer 61

RESULT 40
US-08-237-401A-31
Sequence 31, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-237-401A-31

Alignment Scores:
Pred. No.: 474 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-237-401A-31 (1-68)

QY 1175 GGAGTGTTCAGACTTCAGT 1195
DB 55 GlyValPheThrSerSer 61

RESULT 41
US-09-134-001C-5401
Sequence 5401, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5401
LENGTH: 69
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5401

Alignment Scores:
Pred. No.: 473 Length: 69
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-134-001C-5401 (1-69)

QY 963 AACACTGTGAAGTACTC 983
DB 9 AcnThrLeuGluLysLeuLeu 15

RESULT 42
US-09-107-532A-5492
Sequence 5492, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts

COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5492:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...73
SEQUENCE DESCRIPTION: SEQ ID NO: 5492:
US-09-107-532A-5492

Alignment Scores:
Pred. No.: 470 Length: 73
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-5492 (1-73)
OY 2263 GGCTTAGAGCCATCTCCATC 2243
Db 51 G1yLeuArGa1a1eSer1le 57

RESULT 43
US-08-469-537A-47
Sequence 47, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisondier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: US/08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US/07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-47

Alignment Scores:
Pred. No.: 467 Length: 77
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-469-537A-47 (1-77)
OY 1175 GGAGTGTTCAGACTTCCACT 1195
Db 57 G1yAlpheThr1rSer 63

RESULT 44
US-08-469-537A-48
Sequence 48, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisondier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: US/08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US/07/736,559
FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-48

Alignment Scores:
Pred. No.: 467
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 2
Matches: 77
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-469-537A-48 (1-77)

QY 1175 GGAGGTTCGACCTTCACCT 1195
Db 57 GYVAlpHeThrThrSerSer 63

RESULT 45
US-09-014-969-32
Sequence 32, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-014-969-32

Alignment Scores:
Pred. No.: 464
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 2
Matches: 81
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-014-969-32 (1-81)

QY 672 TCCAGCTTAAATATTCCTC 652
Db 59 SerSerLeuLysLeuSer 65

RESULT 46
US-09-732-210-270
Sequence 270, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitnick, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 270
LENGTH: 88
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-732-210-270

Alignment Scores:
Pred. No.: 460
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 4
Matches: 88
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-732-210-270 (1-88)

QY 2406 CTTGCCGACGTTGTTTCG 2386
Db 45 LeuCyArgArgCysPheArg 51

RESULT 47
US-09-252-991A-19013
Sequence 19013, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19013
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19013

Alignment Scores:
Pred. No.: 458 Length: 90
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-19013 (1-90)

OY 1635 TCTTCAGACGAGACCTCTTCT 1615

Db 20 SerSerSerArgThrSerSer 26

RESULT 48

US-09-370-838-211
; Sequence 211, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 211
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-211

Alignment Scores:

Pred. No.: 457 Length: 92
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-370-838-211 (1-92)

OY 722 CTGCTGACGACATCCCTTTA 742

Db 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 49
US-09-252-991A-28763
; Sequence 28763, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28763
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28763

Alignment Scores:
Pred. No.: 454 Length: 98
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-28763 (1-98)

OY 284 CATTTCTTCCCATTTCTGAT 264

Db 78 HicYSerSerHisSerAsp 84

RESULT 50

US-09-107-532A-5792
; Sequence 5792, Application US/09107532A
; Patent No. 6583275

; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street

CITY: Walham
STATE: Massachusetts
COUNTRY: USA

ZIP: 02354
COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC

OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5792:

SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES

ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...101
SEQUENCE DESCRIPTION: SEQ ID NO: 5792:

US-09-107-532A-5792

Alignment Scores: 452 Length: 101
Pred. No.: 7.00 Matches: 7
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 0.78% Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-107-532A-5792 (1-101)

QY 1423 CACTACATCCAGTCTTCC 1443

Db 14 HieValGlnSerValSer 20

RESULT 51

US-08-905-223-437
; Sequence 437, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 437:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 35..1
; IDENTIFICATION METHOD: Von Heljne matrix
; OTHER INFORMATION: score 7.4
; OTHER INFORMATION: seq SLLSLSLASGSG/CG
US-08-905-223-437

Alignment Scores:

Pred. No.: 451 Length: 103
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 3 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-905-223-437 (1-103)

QY 1248 CTTCTCAGTCACTACGCTA 1228
Db 46 LeuLeuTherSerLeuMetLeu 52

RESULT 52

US-09-107-532A-4380
; Sequence 4380, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 4380:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 111 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...111
; SEQUENCE DESCRIPTION: SEQ ID NO: 4380:
US-09-107-532A-4380

Alignment Scores:

Pred. No.: 447 Length: 111
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-4380 (1-111)

QY 2481 GTCTCTCTGTTCTACCTTC 2461

Db 60 ValLeuLeuPheLeuProPhe 66

RESULT 53

US-09-107-532A-3724
; Sequence 3724, Application US/09107532A

Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3724:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...119
SEQUENCE DESCRIPTION: SEQ ID NO: 3724:
US-09-107-532A-3724

Alignment Scores:
Pred. No.: 443 Length: 119
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-3724 (1-119)

QY 1584 ACTACTTTAAGACGCTCCA 1564
Db 28 ThrThreuaKgrThrValPro 34

RESULT 54
US-09-205-258-268
Sequence 268, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 268
LENGTH: 131
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (131)
OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-268

Alignment Scores:
Pred. No.: 438 Length: 131
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-205-258-268 (1-131)

QY 2520 GCAGCTCAGACACCTCGGCT 2500
Db 122 AlalalacInthreuleugly 128

RESULT 55
US-09-107-532A-5741
Sequence 5741, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Walcham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denek
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5741:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...131
SEQUENCE DESCRIPTION: SEQ ID NO: 5741:
US-09-107-532A-5741

Alignment Scores:
Pred. No.: 438 Length: 131
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-5741 (1-131)

QY 30 TTAGTATCGATGCGACGTCG 50
Db 18 Leuvalillephalathrer 24

RESULT 56
US-09-252-991A-25510
Sequence 25510, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25510
LENGTH: 137
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25510

Alignment Scores:
Pred. No.: 435 Length: 137
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-25510 (1-137)

QY 1328 CGGCGAGAACTTGCCGACG 1348
Db 50 ArgptroglYthrcyproser 56

RESULT 57
US-09-120-663-4
Sequence 4, Application US/09120663
Patent No. 6228644
GENERAL INFORMATION:
APPLICANT: Bogdanove, Adam J.
APPLICANT: Kim, Jihyun Francis
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM
TITLE OF INVENTION: ERWINIA AMYLOVORA, ITS USE, AND ENCODING GENE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/120,663
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/055,105
FILING DATE: 06-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1661
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-120-663-4

Alignment Scores:
Pred. No.: 435 Length: 139
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-120-663-4 (1-139)

QY 272 CATCTGATAGTTCCTTGG 252
Db 48 HisserapSerLeuLeu 54

RESULT 58
5169835-18
PATENT NO. 5169835
APPLICANT: WAI-YEE, CHAN
TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEINS APPLICATIONS
NUMBER OF SEQUENCES: 48
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,409
FILING DATE: 07-AUG-1989
SEQ ID NO: 18
LENGTH: 143
5169835-18

Alignment Scores:
Pred. No.: 433 Length: 143
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x 5169835-18 (1-143)

QY 722 CTGCTGACGACATCCCTTTTA 742
|||||

Db 19 LeuLeuThralaseLeuLeu 25

RESULT 59
US-09-252-991A-18336
Sequence 18336, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18336
LENGTH: 148
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18336

Alignment Scores:
Pred. No.: 431 Length: 148
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-18336 (1-148)

QY 873 AGTTCATTAGTCCACCTCA 893
Db 3 SerSerLeuValProThrSer 9

RESULT 60
US-09-252-991A-31728
Sequence 31728, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31728
LENGTH: 150
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31728

Alignment Scores:
Pred. No.: 431 Length: 150
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-31728 (1-150)

QY 2089 AGATCAAGCCTTCTGCTACG 2109
Db 92 ArgSerArgProSerAlaThr 98
|||||

RESULT 61.
US-09-053-197A-19
; Sequence 19, Application US/09053197A
; Patent No. 6023952
; GENERAL INFORMATION:
; APPLICANT: Turner, Joel H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,197A
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Macknight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UALB-03293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-053-197A-19
Alignment Scores:
Pred. No.: 428 Length: 158
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 3
US-09-774-490-1 (1-2709) x US-09-053-197A-19 (1-158)
QY 2207 ACAGAGCATTTGGAAGACTT 2227
Db 1 ThrluHlsleuGlulguLeu 7
RESULT 62
US-09-085-761A-19
; Sequence 19, Application US/09085761A
; Patent No. 6335178
; GENERAL INFORMATION:
; APPLICANT: Weiner, Joel H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America

ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,761A
; FILING DATE: 28-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03356
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-085-761A-19
Alignment Scores:
Pred. No.: 428 Length: 158
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 4
US-09-774-490-1 (1-2709) x US-09-085-761A-19 (1-158)
QY 2207 ACAGAGCATTTGGAAGACTT 2227
Db 1 ThrluHlsleuGlulguLeu 7
RESULT 63
US-09-134-001C-5575
; Sequence 5575, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lytn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5575
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5575
Alignment Scores:
Pred. No.: 428 Length: 159
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4
US-09-774-490-1 (1-2709) x US-09-134-001C-5575 (1-159)
QY 280 GTTCTCCCATTCGTGATAGTT 260

Db 129 ValLeuProLeuLeuVal 135

RESULT 64
US-09-252-991A-31464
Sequence 31464, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31464
LENGTH: 161
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31464

Alignment Scores:
Pred. No.: 427 Length: 161
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-31464 (1-161)

QY 1527 TCTTCTGCATCCTCGCTCT 1507
Db 56 SerSerAlaSerThrArgSer 62

RESULT 65
US-09-252-991A-20901
Sequence 20901, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20901
LENGTH: 162
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20901

Alignment Scores:
Pred. No.: 427 Length: 162
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-20901 (1-162)

QY 1635 TCTTCAGACAGAACTCTTCT 1615
|||||

Db 134 SerSerSerArgThrSerSer 140

RESULT 66
US-08-311-731A-279
Sequence 279, Application US/08311731A
Patent No. 6583266

GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 279:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Mycobacterium leprae
US-08-311-731A-279

Alignment Scores:
Pred. No.: 426 Length: 165
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-311-731A-279 (1-165)

QY 2308 GTACCAAGCTTCTGCTAGG 2288
Db 143 ValProAspLeuLeuAlaArg 149

RESULT 67
US-09-252-991A-19472
Sequence 19472, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19472
LENGTH: 166
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19472

Alignment Scores:
Pred. No.: 425 Length: 166
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-19472 (1-166)

QY 966 ACTGTGAAAGCTACTCAG 986
Db 105 Thlredujnylsleuthr 111

RESULT 68

US-08-469-537A-86
Sequence 86, Application US/08469537A
Patent No. 5843749

GENERAL INFORMATION:
APPLICANT: Maisompierre, et al.
TITLE OF INVENTION: EHK AND FOR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721

TELEX:
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:

LENGTH: 172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-86

Alignment Scores:
Pred. No.: 423 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-08-469-537A-86 (1-172)

QY 1175 GAGGTGTTTACGACTTCAGT 1195
Db 92 Glyvalpherthrthrserser 98

RESULT 69

US-09-252-991A-18791
Sequence 18791, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18791
LENGTH: 172
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18791

Alignment Scores:

Pred. No.: 423 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-18791 (1-172)

QY 1635 TCTTCAGCAGAACCTCTTCT 1615
Db 76 SerSerSerThrThrSerSer 82

RESULT 70

US-09-252-991A-24167
Sequence 24167, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24167
LENGTH: 177
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24167

Alignment Scores:

Pred. No.: 422 Length: 177
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-252-991A-24167 (1-177)

QY 2118 TGTAGACTACGACGACGACG 2098
DB 153 CysArgLeuThrAlaArgAla 159

RESULT 71
US-09-335-409-14
Sequence 14, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-335-409-14 (1-184)

QY 239 GTATTACTTACGACGACGACG 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 72
US-09-568-102-14
Sequence 14, Application US/09568102
Patent No. 6346404
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum

US-09-568-102-14
Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-568-102-14 (1-184)

QY 239 GTATTACTTACGACGACGACG 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 73
US-09-567-969-14
Sequence 14, Application US/09567969
Patent No. 6355457
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,969
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-567-969-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-567-969-14 (1-184)

QY 239 GTATTACTTACGACGACGACG 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 74
US-09-568-480-14
Sequence 14, Application US/09568480
Patent No. 6353458
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,480
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID*NO 14
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-14

```

```

Alignment Scores:
Pred. No.: 420          Length: 184
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 4                  Gaps: 0

```

US-09-774-490-1 (1-2709) x US-09-568-486-14 (1-184)

```

QY      239 GTATTACTTACAGCAGAGCA 259
      |||||||
Db      58 ValleuLeuthrAlaAArgAla 64

```

RESULT 75

```

US-09-568-486-14
; Sequence 14, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligou, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPROTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-486-14

```

```

Alignment Scores:
Pred. No.: 420          Length: 184
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 4                  Gaps: 0

```

US-09-774-490-1 (1-2709) x US-09-568-486-14 (1-184)

```

QY      239 GTATTACTTACAGCAGAGCA 259
      |||||||
Db      58 ValleuLeuthrAlaAArgAla 64

```

Search completed: August 3, 2003, 10:55:30
Job time : 74.5 secg

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 17:20:41 ; Search time 9580 Seconds
(without alignments)
11568.286 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttaccgatg.....agcttttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : GenEmbl:

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2709	100.0	2709	6 AX207154	AX207154 Sequence
2	2508	92.6	2530	6 AX743092	AX743092 Sequence
3	2508	92.6	2530	6 BD084822	BD084822 Diagnosis
4	2508	92.6	2530	6 HUMHEM	L26081 Homo sapien
5	2508	92.6	2530	11 G31703	G31703 SMS1973 Ex
6	2508	92.6	2601	6 AR040598	AR040598 Sequence
7	2508	92.6	2601	6 I47054	I47054 Sequence
8	1253	46.3	1481	6 I11900	I11900 Sequence 1
9	652	24.1	169408	2 AC073110	AC073110 Homo sapi
10	601	22.2	179640	9 AC006322	AC006322 Homo sapi
11	303	11.2	354	6 AX193169	AX193169 Sequence
12	220	8.1	94027	9 AC073221	AC073221 Homo sapi
13	194	7.2	108642	9 AC004451	AC004451 Homo sapi
14	160	5.9	135214	9 AC004848	AC004848 Homo sapi
15	148	5.5	169408	2 AC073110	AC073110 Homo sapi
16	117	4.3	81667	2 AC091270	AC091270 Homo sapi
17	71	2.6	2137	6 AX207159	AX207159 Sequence
18	68	2.5	2319	10 RNSIICN1	X95286 R.norvegicu
19	63	2.3	229	6 BD026302	BD026302 Sequence
20	63	2.3	110000	2 AC109862.2	Continuation (3 of
21	55	2.0	81667	2 AC091270	AC091270 Homo sapi
22	50	1.8	179667	2 AC095790	AC095790 Rattus no
23	50	1.8	273260	2 AC123888	AC123888 Rattus no
24	47	1.7	5952	10 MUSC1	D85028 Mus musculu
25	47	1.7	195335	2 AC068065	AC068065 Mus muscu
26	47	1.7	265982	2 AC121125	AC121125 Mus muscu
27	41	1.5	1998	10 MUSEMALI1	I40484 Mus musculu
28	41	1.5	2319	10 MUSCOSE	L41541 Mus musculu
29	41	1.5	2913	10 MMRNASEMD	X85993 M.musculu
30	41	1.5	110000	2 AC109862.1	Continuation (2 of
31	41	1.5	184006	2 AC102452	AC102452 Mus muscu
32	41	1.5	207757	10 AC022368	AC022368 Mus muscu
33	33	1.2	195335	2 AC068065	AC068065 Mus muscu
34	33	1.2	195311	10 AC121585	AC121585 Mus muscu
35	31	1.1	4661	6 AX359936	AX359936 Sequence
36	31	1.1	4661	6 AX382150	AX382150 Sequence
37	31	1.1	6620	12 SYNMLPLN3	M28247 Retroviral
38	31	1.1	6837	12 SYNMLP2	M64754 Moloney mur
39	31	1.1	7160	6 AR302096	AR302096 Sequence
40	31	1.1	7235	6 AR302097	AR302097 Sequence
41	29	1.1	29	6 AX591134	AX591134 Sequence
42	27	1.0	3148	5 AF086761	AF086761 Danto rer
43	27	1.0	9729	6 AX107928	AX107928 Sequence
44	27	1.0	12473	6 AX107927	AX107927 Sequence
45	25	0.9	25	6 AX591129	AX591129 Sequence
46	25	0.9	1400	6 AX207158	AX207158 Sequence
47	25	0.9	244749	2 AC098444	AC098444 Rattus no
48	25	0.9	250672	2 AC096525	AC096525 Rattus no
49	24	0.9	24	6 AX591128	AX591128 Sequence
50	24	0.9	3263	5 GGU02528	GU02528 Gallus galli
51	24	0.9	225005	2 AC102985	AC102985 Rattus no
52	23	0.8	29	6 AX591132	AX591132 Sequence
53	23	0.8	1822	10 D21099	D21099 Mouse mRNA
54	23	0.8	2325	5 AY030051	AY030051 Xenopus 1
55	23	0.8	51312	2 AC100365	AC100365 Mus muscu
56	23	0.8	225372	2 AC120544	AC120544 Mus muscu
57	22	0.8	883	5 GGU02841	GU02841 Gallus galli
58	22	0.8	5732	6 AX359932	AX359932 Sequence
59	22	0.8	5732	6 AX382146	AX382146 Sequence
60	22	0.8	7311	6 AB0476	AB0476 Sequence 3
61	22	0.8	7311	6 BD136840	BD136840 Transoma
62	22	0.8	7352	6 AR302095	AR302095 Sequence
63	22	0.8	7353	6 AR302092	AR302092 Sequence
64	22	0.8	7591	6 AX107929	AX107929 Sequence
65	22	0.8	7885	6 AB0477	AB0477 Sequence 4

Pred. No. is the number of results predicted by chance to have a

c	66	22	0.8	7885	6	BD136841	BD136841 Transsoma
c	67	22	0.8	9183	6	AX359933	AX359933 Sequence
c	68	22	0.8	9183	6	AX382147	AX382147 Sequence
c	69	22	0.8	90604	9	AC006525	AC006525 Homo sapi
c	70	22	0.8	110000	2	AC112426_0	AC112426 Rattus no
c	71	22	0.8	110000	2	AC112426_2	Continuation (3 of
c	72	22	0.8	150638	2	AC129994	AC129994 Rattus no
c	73	22	0.8	166965	2	AC122712	AC122712 Homo sapi
c	74	22	0.8	216124	2	AC098305	AC098305 Rattus no
c	75	22	0.8	222086	2	AC106106	AC106106 Rattus no
c	76	22	0.8	276158	2	AC129229	AC129229 Rattus no
c	77	22	0.8	282515	2	AC118094	AC118094 Rattus no
c	78	22	0.8	298698	2	AC094778	AC094778 Rattus no
c	79	22	0.8	328905	2	AC094646	AC094646 Rattus no
c	80	21	0.8	591133	29	AC591133	Sequence
c	81	21	0.8	618	11	BV056652	BV056652 S21P6751
c	82	21	0.8	1385	6	AX174768	AX174768 Sequence
c	83	21	0.8	1385	6	BD087170	BD087170 Method of
c	84	21	0.8	1661	6	AR217930	AR217930 Sequence
c	85	21	0.8	1661	6	BD057812	BD057812 Growth fa
c	86	21	0.8	1808	6	AX283629	AX283629 Sequence
c	87	21	0.8	2026	6	AR179702	AR179702 Sequence
c	88	21	0.8	3091	6	AX139124	AX139124 Sequence
c	89	21	0.8	4207	6	AX359930	AX359930 Sequence
c	90	21	0.8	4207	6	AX382144	AX382144 Sequence
c	91	21	0.8	4210	6	AX359931	AX359931 Sequence
c	92	21	0.8	4210	6	AX382145	AX382145 Sequence
c	93	21	0.8	6255	6	AX359939	AX359939 Sequence
c	94	21	0.8	6255	6	AX382153	AX382153 Sequence
c	95	21	0.8	9511	6	AX359960	AX359960 Sequence
c	96	21	0.8	9511	6	AX382174	AX382174 Sequence
c	97	21	0.8	37672	6	AX695374	AX695374 Sequence
c	98	21	0.8	137942	10	AL645745	AL645745 Mouse DNA
c	99	21	0.8	163245	2	AC013699	AC013699 Homo sapi
c	100	21	0.8	165940	2	AC109753	AC109753 Rattus no

ALIGNMENTS

RESULT 1
AX207154
LOCUS
DEFINITION
AX207154
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN

AX207154
Sequence 1 from Patent WO0155455.
AX207154.1 GI:15394943
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Jin, S.
Resistance sequences and uses thereof
Patent: WO 015545-A 1 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US); Jin, Shengfang (US)
location/Qualifiers
1..2709
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

831 a 560 c 620 g 698 t

Query Match 100.0%; Score 2709; DB 6; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGATTAACAAGCTTAGTATGATGACGACGAGGAGTGTG 60
DB 1 AATCTTTATTTATGATGATTAACAAGCTTAGTATGATGACGACGAGGAGTGTG 60
QY 61 ACCCAGCGCTCGGAGTAGTGAAGCTCGCTGTTCTCCATTTGTACAGCAGTATTT 120

DB 61 ACCCAGCGCTCGGAGTAGTGAAGCTCGCTGTTCTCCATTTGTACAGCAGTATTT 120
QY 121 CGAGATTTGTTAACTTCTCGGCCGCAAAATACAGAAAGAAAGCTAAAGAGCAAGG 180
DB 121 CGAGATTTGTTAACTTCTCGGCCGCAAAATACAGAAAGAAAGCTAAAGAGCAAGG 180
QY 181 GACCTACAGCGCTCGCAGCATGGGCTGGTAACTAGATTTGTCCTCTTTCTGGGAGT 240
DB 181 GACCTACAGCGCTCGCAGCATGGGCTGGTAACTAGATTTGTCCTCTTTCTGGGAGT 240
QY 241 ATTACTTACAGCAAGCAAGCAACTATCAGATGGGAAACAAATGTCACAGGCTGAATT 300
DB 241 ATTACTTACAGCAAGCAAGCAACTATCAGATGGGAAACAAATGTCACAGGCTGAATT 300
QY 301 ATCCCTACAAAGAAAGTTGGAATCCAAAGTGTATCTTTCAATGGCTTGCCAAACG 360
DB 301 ATCCCTACAAAGAAAGTTGGAATCCAAAGTGTATCTTTCAATGGCTTGCCAAACG 360
QY 361 CTCAGTTATCATCTTCTTTGATGAGGAACGAGTAGGCTGTATGTTGAGCAAA 420
DB 361 CTCAGTTATCATCTTCTTTGATGAGGAACGAGTAGGCTGTATGTTGAGCAAA 420
QY 421 GGATCATATTTTATTTGACCTTGCTTAATATCAGGATTTTCAAAGATTGTGTGCC 480
DB 421 GGATCATATTTTATTTGACCTTGCTTAATATCAGGATTTTCAAAGATTGTGTGCC 480
QY 481 AGTATCTTAACACAGAGAGTGAATGCAAGGGCTGAAAAGACATCTGAAAGATG 540
DB 481 AGTATCTTAACACAGAGAGTGAATGCAAGGGCTGAAAAGACATCTGAAAGATG 540
QY 541 TGTCTAATTTCAATCAAGTACTTAAGGCAATATATCAGCTCATTTGACGCTGTGAA 600
DB 541 TGTCTAATTTCAATCAAGTACTTAAGGCAATATATCAGCTCATTTGACGCTGTGAA 600
QY 601 GGGGGCTTTTCAATTCATTTGCACTTGAATTTGCAATCTCTGAGAGCAATAT 660
DB 601 GGGGGCTTTTCAATTCATTTGCACTTGAATTTGCAATCTCTGAGAGCAATAT 660
QY 661 TTTTAAAGTGGAGAACTACATTTTGAAGAAAGGGCGTGGAGAGTCCATATGACCTTA 720
DB 661 TTTTAAAGTGGAGAACTACATTTTGAAGAAAGGGCGTGGAGAGTCCATATGACCTTA 720
QY 721 GCTGCTGACAGCATCTTTTAAATAGATGAGAAATTAATCTCTGGAAGCTGAGATTT 780
DB 721 GCTGCTGACAGCATCTTTTAAATAGATGAGAAATTAATCTCTGGAAGCTGAGATTT 780
QY 781 TATGGGGGAGAACTTTGCTATCTTCCGAATCTTGGGACCAACCAATCAGAGAGA 840
DB 781 TATGGGGGAGAACTTTGCTATCTTCCGAATCTTGGGACCAACCAATCAGAGAGA 840
QY 841 GCAGCATATTCAGAGTGGCTCAATGATCAAAAGTTCAATAGTCCACCTCATCTCAGA 900
DB 841 GCAGCATATTCAGAGTGGCTCAATGATCAAAAGTTCAATAGTCCACCTCATCTCAGA 900
QY 901 GAGTACATCTCTGAAGATGAACAAGTACTTTTCTTCGCTGAAAAATGCAATAGATG 960
DB 901 GAGTACATCTCTGAAGATGAACAAGTACTTTTCTTCGCTGAAAAATGCAATAGATG 960
QY 961 AGAATCTCTGAAAAAGCTACACGCTAGAAATAGTCAATATGCAAGATGACTTTGG 1020
DB 961 AGAATCTCTGAAAAAGCTACACGCTAGAAATAGTCAATATGCAAGATGACTTTGG 1020
QY 1021 AGGGCAGAGAGTCTGGATTAATGGAACAATTCCTCAAGCTCGCTGATTTGGCTC 1080
DB 1021 AGGGCAGAGAGTCTGGATTAATGGAACAATTCCTCAAGCTCGCTGATTTGGCTC 1080
QY 1081 AGTGCAGAGTCAATAGCATTTGACATCTATTTGATGACGAGATGATCTTAAT 1140
DB 1081 AGTGCAGAGTCAATAGCATTTGACATCTATTTGATGACGAGATGATCTTAAT 1140
QY 1141 GAATTTAAAGATCTTAAAAATCCAGTTGATATGAGATGTTTACGACTTCAGTAACAT 1200

Dp	1141	GAACCTTTAAAGATCCCTAAAATTCAGGTTGTAATAGAGGTTTACGACTTCCAGTAACAT	1201
Qy	1201	TTTCAAGGGATACAGCCGTGTGTATGTATAGCATGTAGTGTAGAGAGGGTGTCTTTGG	1266
Dp	1201	TTTCAAGGGATACAGCCGTGTGTATGTATAGCATGTAGTGTAGAGAGGGTGTCTTTGG	1266
Qy	1261	TCATATGCCCAAGGGATAGAGCCCACTATCAATGAAGGTGCTTATCAAGGAAGAGTCCC	1320
Dp	1261	TCATATGCCCAAGGGATAGAGCCCACTATCAATGAAGGTGCTTATCAAGGAAGAGTCCC	1320
Qy	1321	CTATCAGGGCCAGAGACTTGTCCAGCAAAAACATTTGGTGGTTTGTACTCTACAAAGGA	1380
Dp	1321	CTATCAGGGCCAGAGACTTGTCCAGCAAAAACATTTGGTGGTTTGTACTCTACAAAGGA	1380
Qy	1381	CCTTCTCTATGATGTTTATPACCTTTGGCAGAAAGTCATCCAGCCATGTACAAATCCAGTGT	1440
Dp	1381	CCTTCTCTATGATGTTTATPACCTTTGGCAGAAAGTCATCCAGCCATGTACAAATCCAGTGT	1440
Qy	1441	TCCTATGAAACAATGCCCCCAATAGATCTCAAAACGGATGTAAATTTATCAATTTACAAAT	1500
Dp	1441	TCCTATGAAACAATGCCCCCAATAGATCTCAAAACGGATGTAAATTTATCAATTTACAAAT	1500
Qy	1501	TGTGCTAAGCCAGATGGAATGACAGAAAGATGACAAGATATGTTATGTTATCGGAACGA	1560
Dp	1501	TGTGCTAAGCCAGATGGAATGACAGAAAGATGACAAGATATGTTATGTTATCGGAACGA	1560
Qy	1561	TGTTGGAGCCGTTCTTAAAGTAGTTCAATTCCTAAGAGACCTTGTATGATTTAGAGA	1620
Dp	1561	TGTTGGAGCCGTTCTTAAAGTAGTTCAATTCCTAAGAGACCTTGTATGATTTAGAGA	1620
Qy	1621	GGTTCTGTGGAAGAAATGACAGTTTTCGGGAACCGACTCTATTCAGCAATGAGACT	1680
Dp	1621	GGTTCTGTGGAAGAAATGACAGTTTTCGGGAACCGACTCTATTCAGCAATGAGACT	1680
Qy	1681	TTCCACTAAGCAGCAACAATATATATGTTCAACCGCTGAGGGTGGCCGAGTCCCTTT	1740
Dp	1681	TTCCACTAAGCAGCAACAATATATATGTTCAACCGCTGAGGGTGGCCGAGTCCCTTT	1740
Qy	1741	ACACCGGTGTATATTTACGGGAAAGCGTGTGCTAGTGTGCTCGCCGAGACCCCTTA	1800
Dp	1741	ACACCGGTGTATATTTACGGGAAAGCGTGTGCTAGTGTGCTCGCCGAGACCCCTTA	1800
Qy	1801	CTGTGCTTGGATGTTCTGCATGTTCTGTCTATTTCCACTGCAAAAGACGCAAG	1860
Dp	1801	CTGTGCTTGGATGTTCTGCATGTTCTGTCTATTTCCACTGCAAAAGACGCAAG	1860
Qy	1861	ACGACAAATATTAAGAAATGAGAACCCACTGACTCACTGTACAGCTTACCAATGATTA	1920
Dp	1861	ACGACAAATATTAAGAAATGAGAACCCACTGACTCACTGTACAGCTTACCAATGATTA	1920
Qy	1921	TCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTGTAGAAATATATACACATT	1980
Dp	1921	TCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTGTAGAAATATATACACATT	1980
Qy	1981	TTTGGATGCAGTCCGAAGTCGACAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAA	2040
Dp	1981	TTTGGATGCAGTCCGAAGTCGACAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAA	2040
Qy	2041	TGAAGAGGGAAGAAAGAGATCAAGTGTGAATCATATCATTCAGAGACAGTCAAGGCT	2100
Dp	2041	TGAAGAGGGAAGAAAGAGATCAAGTGTGAATCATATCATTCAGAGACAGTCAAGGCT	2100
Qy	2101	TCTGTACTGTAGTCTTACAAACGAAGAGATTCAAGCAATTAACCTCTGTCCATGTGGTGAACA	2160
Dp	2101	TCTGTACTGTAGTCTTACAAACGAAGAGATTCAAGCAATTAACCTCTGTCCATGTGGTGAACA	2160
Qy	2161	TGGGTTCAATACAACTCTTTTAAAGTAAACCTGTGAAGTCAATTGACACAGAGCAATTTGA	2220
Dp	2161	TGGGTTCAATACAACTCTTTTAAAGTAAACCTGTGAAGTCAATTGACACAGAGCAATTTGA	2220
Qy	2221	AGAACTTCTTCAATAAGATGATGTAGAGATGGCTTAAAGCAAAAGAAATGTCCAATAG	2280
Dp	2221	AGAACTTCTTCAATAAGATGATGTAGAGATGGCTTAAAGCAAAAGAAATGTCCAATAG	2280

QY	2281	CATGACACCTTAGCCAGAAAGTCTGTGATCAGAGACTTCAATGCACTCATCAACCCCA	2340
Db	2281	CATGACACCTTAGCCAGAAAGTCTGTGATCAGAGACTTCAATGCACTCATCAACCCCA	2340
QY	2341	TCTCAACGACGATGAGTGTCTGTGACCAAGTTTGAAAGGACCGAACAACGTCG	2400
Db	2341	TCTCAACGACGATGAGTGTCTGTGACCAAGTTTGAAAGGACCGAACAACGTCG	2400
QY	2401	GCAAAAGCCAGAGACATACCCAGGGAAACAGTAAACAATGTAAGCACTTACAAGAAATTA	2460
Db	2401	GCAAAAGCCAGAGACATACCCAGGGAAACAGTAAACAATGTAAGCACTTACAAGAAATTA	2460
QY	2461	GAAAGTGAAGAAACAGAGAGACCCAGAAATTTGAAGGACCCAGAGAGTGTCTGAGCTGC	2520
Db	2461	GAAAGTGAAGAAACAGAGAGACCCAGAAATTTGAAGGACCCAGAGAGTGTCTGAGCTGC	2520
QY	2521	ATTACCTTAAAGAACTCTAAACAAGTAAACAATTGCCCTAGACATTAATGGAAGAAACAA	2580
Db	2521	ATTACCTTAAAGAACTCTAAACAAGTAAACAATTGCCCTAGACATTAATGGAAGAAACAA	2580
QY	2581	TGCATATATACATGAACCTTTTTCATATGCGATATGATGTGATTTCAATGATGGGAAATTC	2640
Db	2581	TGCATATATACATGAACCTTTTTCATATGCGATATGATGTGATTTCAATGATGGGAAATTC	2640
QY	2641	AGCTGAGTTCACCAATTAATAATTAATTAATTCATGAGTAACCTTCTTAATAGCTTTT	2700
Db	2641	AGCTGAGTTCACCAATTAATAATTAATTAATTCATGAGTAACCTTCTTAATAGCTTTT	2700
QY	2701	CCTAATACC 2709	
Db	2701	CCTAATACC 2709	
RESULT 2			
LOCUS	AX743092	2530 bp	DNA linear PAT 12-MAY-2003
DEFINITION	Sequence 9 from Patent WO03029814.		
ACCESSION	AX743092		
VERSION	AX743092.1	GI:30577056	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
JOURNAL	Altalo, K., Karikainen, M. and Katila, K.		
FEATURES	Neutrophilin/vegf c/vegf 3 matrisale and methods		
Source	Patent: WO 03029814-A 9 10-Apr-2003;		
	LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; Licentia Ltd. (FI)		
	Location/Qualifiers		
	1..2530		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	16..2331		
	/codon_start=1		
	/protein_id="CAD91050.1"		
	/db_xref="GI:30577057"		
	/translation="MGWVTRIVCLFMGVLLTARANYONGKNVPRKLSTYKEMLENNNN		
	VIYFNGLANSSVHTPLDEERSRLVYGAADHIFSPDLVNIKDFQIWPVSYTRDDEI		
	CKMAGKDIIEKCANPIKLVKAYNQTHTVAGSTGAFFHICVYIEIGHHPENDI		
	HEFNGRKSPIYDPLKLRSLIDBELYSGTAPDMGDDPAIFRLTGHNNRIRTDHDS		
	RMLNDPFIISHLISDSNPDDKVIYFERNALIDGHSAGTHARIGQICNDPFGH		
	RSLVNKPFTFLKARLICSVPNGIDHPELODFLNNKEDPNPVVYGFVTSNVI		
	FKSGAVMSMSDVRVFLGYPARHDPNQPWYQGRVYPRGTSKTFGFDST		
	KDLDDVITTFARSHAPMTNVPFEMNNRPIVKTQVNYQFOIVVDRVDAEDGQYDME		
	IGDVPYLRCDIYGAACAECCLARPYCAMGSACSRFPFPAKRRTRQDRNDPLTH		
	CSDLAHNNHGHSPFEERIITGVENSSTFLSCSPSQALVYWGQQRNERKEIRVD		
	DHIIITQGLLRSLQKDGSAVLCVAHEHGFLOTLLKVLTEVIDTEHLELLHKDD		
	GDGSKTKEMSNMTSPQKWYRNPFWMLINHPNINTNDEPCEQWKRDKRQRQPGHT		

PENSNKXHLQENKGRNRTHFERAPRSV"
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN
Query Match 92.6%; Score 2508; DB 6; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 CTGCAGCATGGGCTGGTAACTAGATTGCTCTCTTTCTGGGAGTATTAATTACACG 252
DB 9 CTGCAGCATGGGCTGGTAACTAGATTGCTCTCTTTCTGGGAGTATTAATTACACG 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAGAACAATGTGCCAAGGCTGAAATTAATCTCAAAAGA 312
DB 69 AAGAGCAAACTATCAGAAATGGGAAGAACAATGTGCCAAGGCTGAAATTAATCTCAAAAGA 128
QY 313 AATGTGGAAATCCAAATGTATGTCATTTCATATGCTTGGCCAAAGCTCCAGTTATCA 372
DB 129 AATGTGGAAATCCAAATGTATGTCATTTCATATGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACAAAGATCAATATT 432
DB 189 TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACAAAGATCAATATT 248
QY 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 492
DB 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 308
QY 493 CAGAAGATGATGCAAGTGGGCTGGAAAGACATCCGAAAGATGTATATTTTCAAT 552
DB 309 CAGAAGATGATGCAAGTGGGCTGGAAAGACATCCGAAAGATGTATATTTTCAAT 368
QY 553 CAAGGACTTAAAGGCATATATATCAGACTCACTGTATGAGCTGTGAAACGGGGCTTTTGA 612
DB 369 CAAGGACTTAAAGGCATATATATCAGACTCACTGTATGAGCTGTGAAACGGGGCTTTTGA 428
QY 613 TCCAAATTTGCACTTACATTTGAATTTGAATCTCTGAGGACAAATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACTTACATTTGAATTTGAATCTCTGAGGACAAATATTTTAAAGCTGGA 488
QY 673 GAACCTCAATTTTGAAGACGGCGGTGGGAAGGTCCATATGACCCATACCTGTGACAC 732
DB 489 GAACCTCAATTTTGAAGACGGCGGTGGGAAGGTCCATATGACCCATACCTGTGACAC 548
QY 733 ATCCCTTTAATAGATGAGAAATTAATCTCTGGAATGCAAGCTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTAATAGATGAGAAATTAATCTCTGGAATGCAAGCTGATTTTATGGGCGAGA 608
QY 793 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGACAGATGATTC 852
DB 609 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGACAGATGATTC 668
QY 853 CAGGTGGCTCATGATGCCAAGTTCATTTAGTGGCCACTCATCTCAGAGATGCAATCC 912
DB 669 CAGGTGGCTCATGATGCCAAGTTCATTTAGTGGCCACTCATCTCAGAGATGCAATCC 728
QY 913 TGAAGATGCAAAAGTATCTTTTCTTCGCTGAAAATGCAATAGATGAGAACACTCTG 972
DB 729 TGAAGATGCAAAAGTATCTTTTCTTCGCTGAAAATGCAATAGATGAGAACACTCTG 788
QY 973 AAAAGTACTACGCTAGATAGTGCATATGCAAGAAATGACTTTGAGGGACAGANG 1032
DB 789 AAAAGTACTACGCTAGATAGTGCATATGCAAGAAATGACTTTGAGGGACAGANG 848
QY 1033 TCTGTGAATTAATGGAACAATTCCTCAAGTGTCTGATTTGTCTGAGTCCAGSTTC 1092
DB 849 TCTGTGAATTAATGGAACAATTCCTCAAGTGTCTGATTTGTCTGAGTCCAGSTTC 908
QY 1093 AAATGSCATTTGACACTCATTTTGAATGAACTGAGAGATTTCTTAATGAATTTAAAGA 1152
DB 909 AAATGSCATTTGACACTCATTTTGAATGAACTGAGAGATTTCTTAATGAATTTAAAGA 968
QY 1153 TCTTAAAAATCCAGTTGTATATGAGTGTTTACGACTTCCAGTAACTTTTCAAGGATTC 1212

DB 969 TCTTAAAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACTTTTCAAGGATTC 1028
QY 1213 ACCCGTGTATATGATATACATAGATGATGAGAGGGGTTCCTTGGCCATATGCCA 1272
DB 1029 ACCCGTGTATATGATATACATAGATGATGAGAGGGGTTCCTTGGCCATATGCCA 1088
QY 1273 CAGGATGGAACCCCACTATCAATGGGTCCCTTATCAAGGAAGTCCCTTATCCAGGCC 1332
DB 1089 CAGGATGGAACCCCACTATCAATGGGTCCCTTATCAAGGAAGTCCCTTATCCAGGCC 1148
QY 1333 AGGAATCTGTCCAGCAAAAATTTGGTGTTTTACCTTACAAAGAACCTTCTGATGA 1392
DB 1149 AGGAATCTGTCCAGCAAAAATTTGGTGTTTTACCTTACAAAGAACCTTCTGATGA 1208
QY 1393 TGTATAACCTTTGCAAGAGATCATCCAGCATGATGATCAATCCAGTTTCTATGAACA 1452
DB 1209 TGTATAACCTTTGCAAGAGATCATCCAGCATGATGATCAATCCAGTTTCTATGAACA 1268
QY 1453 TCGCCCAATAGTATCAAAAACGATGTAATTTATCAATTTACAAATTTGTCTAGACCG 1512
DB 1269 TCGCCCAATAGTATCAAAAACGATGTAATTTATCAATTTATCAAAATTTGTCTAGACCG 1328
QY 1513 AGTGAATGCAAGATGGAACAGTATGATGTTTATCGGAACAGATTTGGACCGT 1572
DB 1329 AGTGAATGCAAGATGGAACAGTATGATGTTTATCGGAACAGATTTGGACCGT 1388
QY 1573 TCTTAAAGTATTTCAATTTCCATAGAGACTTGTATGATTTGAAGAGGTTCTGCTGA 1632
DB 1389 TCTTAAAGTATTTCAATTTCCATAGAGACTTGTATGATTTGAAGAGGTTCTGCTGA 1448
QY 1633 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTACGAATGAGCTTTCACATAGCA 1692
DB 1449 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTACGAATGAGCTTTCACATAGCA 1508
QY 1693 GCAAACTATATATTTGTTCAAGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGGA 1752
DB 1509 GCAAACTATATATTTGTTCAAGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGGA 1568
QY 1753 TATTTACGGGAAGAGGCTGTGCTGAGTGTGCTGCCCGGAGCCCTTACTGTGCTGGGA 1812
DB 1569 TATTTACGGGAAGAGGCTGTGCTGAGTGTGCTGCCCGGAGCCCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCGATGTTCTGCTATTTTCCCATGTCGAAGAGACGCAACAGATAT 1872
DB 1629 TGGTTCGATGTTCTGCTATTTTCCCATGTCGAAGAGACGCAACAGATAT 1688
QY 1873 AAGAAATGGAACCCCACTGACTCACTGTTCAAGCTTACACCATATATACCATGGCCA 1932
DB 1689 AAGAAATGGAACCCCACTGACTCACTGTTCAAGCTTACACCATATATACCATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGTAGACATTTTGGAAATGAG 1992
DB 1749 CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGTAGACATTTTGGAAATGAG 1808
QY 1993 TCCGAAGTCCGAGAGAGGCTGTGCTATTTGCAATTTCCAGAGGCGAAATGAAGCGAAA 2052
DB 1809 TCCGAAGTCCGAGAGAGGCTGTGCTATTTGCAATTTCCAGAGGCGAAATGAAGCGAAA 1868
QY 2053 AGAAGATCAGAGTGAATGATCATATCATATCAGGACAAATCAAGGCTTTCGCTACGTAG 2112
DB 1869 AGAAGATCAGAGTGAATGATCATATCATATCAGGACAAATCAAGGCTTTCGCTACGTAG 1928
QY 2113 TCTACAGAGAGATTCAGGCAATTTACCTGTGCACTCGGTGGAACATGGGTTATATCA 2172
DB 1929 TCTACAGAGAGATTCAGGCAATTTACCTGTGCACTCGGTGGAACATGGGTTATATCA 1988
QY 2173 AACTCTTTTAAGGTAACCTGTGAAGTCAATGACACAGAGATTTGGAAGAACTTCTTCA 2232
DB 1989 AACTCTTTTAAGGTAACCTGTGAAGTCAATGACACAGAGATTTGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCATATAGATGACACCTAG 2292

Db 2049 TAAAGATGATGATGATGCTCTAAGACCAAGAAATGTCATAGCATGACACTAG 2108
Qy 2293 CCAGAGAGCTGCTGTCAGAGACTTCATGAGCTCATCAACGACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGAGCTGCTGTCAGAGACTTCATGAGCTCATCAACGACCCCAATCTCAACAGAT 2168
Qy 2353 GGATGAGTCTGTGAAACAAGTTGGAAAAAGGACCGAAAAACAAGCTGGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAAACAAGTTGGAAAAAGGACCGAAAAACAAGCTGGCAAGGCCAGG 2228
Qy 2413 ACATCCCCAGGAAACAGTAAACAATGGAAGCACTTACAGAAATTAAGAAAGTAAAGAA 2472
Db 2229 ACATCCCCAGGAAACAGTAAACAATGGAAGCACTTACAGAAATTAAGAAAGTAAAGAA 2288
Qy 2473 CAGAGAGACCCAGAAATTTGAGAGGCAACCCAGAGAGTGTCTGAGCTGCTACTCTGA 2532
Db 2289 CAGAGAGACCCAGAAATTTGAGAGGCAACCCAGAGAGTGTCTGAGCTGCTACTCTGA 2348
Qy 2533 AACCTCAACAGATGAGAACTTGCTCTAGACATTAACCTGAAAAACAATGCAATTAAT 2592
Db 2349 AACCTCAACAGATGAGAACTTGCTCTAGACATTAACCTGAAAAACAATGCAATTAAT 2408
Qy 2593 GAACCTTTTTCATGAGCATTAATGATGATGTTTAAATGTTGGAAATTCAGCTGAGTTCA 2652
Db 2409 GAACCTTTTTCATGAGCATTAATGATGATGTTTAAATGTTGGAAATTCAGCTGAGTTCA 2468
Qy 2653 CCAATTATTAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
Db 2469 CCAATTATTAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2516

RESULT 3
BD084822 2530 bp DNA linear PAT 27-AUG-2002

LOCUS BD084822 2530 bp DNA linear PAT 27-AUG-2002

DEFINITION Diagnosis method and reagents.

ACCESSION BD084822.1 GI:22630432

VERSION BD084822.1

KEYWORDS JP 2001522241-A/15.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2530)

AUTHORS Leeuwen, F. W. V., Grosveld, F. G. and Burbach, J. P. H.

TITLE Diagnosis method and reagents

JOURNAL Patent: JP 2001522241-A 15 13-NOV-2001;
ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCES, ERASMUS UNIVERSITY
ROTTERDAM, UNIVERSITY OF UTRECHT

COMMENT OS Homo sapiens (human)
PN JP 2001522241-A/15
PD 13-NOV-2001
PF 02-APR-1998 JP 1998542545
PR 10-APR-1997 US 60/043163
PI FEDERIK W VAN LEEUWEN, FRANKLIN G GROSVELD, JOHANNES PETER
HENRI BURBACH

PC C120J1/68, C07K14/47, C12N15/52, C12N9/00, C12N5/10, A61K38/43, PC
A01K67/027,
PC A01K48/00//C07K16/18
CC Strandedness: Double;
CC Topology: Linear;
CC semaphorin-III gene, Genbank accession number L26081 FH Key

CC semaphorin-III gene, Genbank accession number L26081 FH Key

FT source 1..2530
location/Qualifiers

FEATURES 1..2530
location/Qualifiers

BASE COUNT 786 a 518 c 576 g 650 t

ORIGIN

Query Match 92.6%; Score 2508; DB 6; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 CTGACGATGGGCTGGTTAACTAGATTTGTCTTTCTTCTGGGAGATTAATCTTACAGC 252
Db 9 CTGACGATGGGCTGGTTAACTAGATTTGTCTTTCTTCTGGGAGATTAATCTTACAGC 68
Qy 253 AAGAGCAATCATGAGATGGAAGAAACAATGTGCAAGGCTGAATATCTCAAGAA 312
Db 69 AAGAGCAATCATGAGATGGAAGAAACAATGTGCAAGGCTGAATATCTTACAAAGA 128
Qy 313 AATGTGGAATCCAAATGTGATCACTTCAATGCTTGGCCAAAGCTCAGTTATCA 372
Db 129 AATGTGGAATCCAAATGTGATCACTTCAATGCTTGGCCAAAGCTCAGTTATCA 188
Qy 373 TACCTTCTTTTGTGATGAGGAAACGAGATGAGCTGTATGTTGAGCAAGAAATCATATT 432
Db 189 TACCTTCTTTTGTGATGAGGAAACGAGATGAGCTGTATGTTGAGCAAGAAATCATATT 248
Qy 433 TTCAATGCACTGGTTAAATCAAGATTTCAAAAGATTTGTGGCCAGTATCTTACAC 492
Db 249 TTCAATGCACTGGTTAAATCAAGATTTCAAAAGATTTGTGGCCAGTATCTTACAC 308
Qy 493 CAGAGAGATGATGCAAGTGGCTGGAAGAAAGATCTGGAAGAAATGTGCTAAATTCAT 552
Db 309 CAGAGAGATGATGCAAGTGGCTGGAAGAAAGATCTGGAAGAAATGTGCTAAATTCAT 368
Qy 553 CAAGTACTTAAGCATTAATCAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 612
Db 369 CAAGTACTTAAGCATTAATCAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 428
Qy 613 TCCAATTTGACCTACATTAATGGAATTTGACATCATCTGAGGACAAATTTTAAAGCTGA 672
Db 429 TCCAATTTGACCTACATTAATGGAATTTGACATCATCTGAGGACAAATTTTAAAGCTGA 488
Qy 673 GAACCTCAATTTGAAAAACGCGGTGGGAAAGTCCATATGACCTTAAGCTGTGACAGC 732
Db 489 GAACCTCAATTTGAAAAACGCGGTGGGAAAGTCCATATGACCTTAAGCTGTGACAGC 548
Qy 733 ATCCCTTTTAATTAATGAGAAATTAATCACTGGAATCTGGAATCTGGAATCTGGAATCT 792
Db 549 ATCCCTTTTAATTAATGAGAAATTAATCACTGGAATCTGGAATCTGGAATCTGGAATCT 608
Qy 793 CTGTGATCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCT 852
Db 609 CTGTGATCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCT 668
Qy 853 CAGGTGCTCAATGATCAAAAGTTCATTAAGTCCACCTCATCTCAGAGATGCAATCC 912
Db 669 CAGGTGCTCAATGATCAAAAGTTCATTAAGTCCACCTCATCTCAGAGATGCAATCC 728
Qy 913 TGAAGATGACAAAGTATCTTTTCTTCCGTAAGAAATGCAATGATGAGAAACCTCTGG 972
Db 729 TGAAGATGACAAAGTATCTTTTCTTCCGTAAGAAATGCAATGATGAGAAACCTCTGG 788
Qy 973 AAAAGCTACTAGCTAGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
Db 789 AAAAGCTACTAGCTAGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 848
Qy 1033 TCTGTGATTAATGAGCAACATCTTCAAAAGCTGCTGATTTTGTCAAGTCCAGGTC 1092
Db 849 TCTGTGATTAATGAGCAACATCTTCAAAAGCTGCTGATTTTGTCAAGTCCAGGTC 908
Qy 1093 AATGTGATGACACTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1152
Db 909 AATGTGATGACACTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 968
Qy 1153 TCCTAAATATCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Db 969 TCCTAAATATCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
Qy 1213 AGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272

Db	1029	AGCGGTGTGTATGTATAGCATGATGATGTGAGAAAGGGGTTCCTTGGTCATATAGCCCA	1088
OY	1273	CAGGGAATGGAACCCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTATCCAGGCGC	1333
Db	1089	CAGGGAATGGAACCCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTATCCAGGCGC	1148
OY	1333	AGGAACCTGTCCCAAGCAAAACATTTGGTGGTTTGACTCACTCAAGAAAGACCTTCCGTATGA	1392
Db	1149	AGGAACCTGTCCCAAGCAAAACATTTGGTGGTTTGACTCACTCAAGAAAGACCTTCCGTATGA	1208
OY	1393	TGTTATTAACCTTTGCAAGAGTGCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA	1453
Db	1209	TGTTATTAACCTTTGCAAGAGTGCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA	1268
OY	1453	TCGCCCAATAGTGAACCAAAAGGATGTAAATTAATCAATTTACAAATTTGTCGTAGACCG	1512
Db	1269	TCGCCCAATAGTGAACCAAAAGGATGTAAATTAATCAATTTACAAATTTGTCGTAGACCG	1328
OY	1513	AGTGAATSCAAGAAATGGAACAGTATGATGTATGTTTATCGGAACAGATGTTGGAGCCGT	1572
Db	1329	AGTGAATSCAAGAAATGGAACAGTATGATGTATGTTTATCGGAACAGATGTTGGAGCCGT	1388
OY	1573	TCTTAAAGTACTTTCATATTCCTTAAGGAACCTTGGATGTATGATTTAAGAAAGGTTCTGCTGGA	1633
Db	1389	TCTTAAAGTACTTTCATATTCCTTAAGGAACCTTGGATGTATGATTTAAGAAAGGTTCTGCTGGA	1448
OY	1633	AGAAATGACAGTTTTTCGGGAAACCACTGCTATTTCAACAAATGAGCTTTCACATACCA	1692
Db	1449	AGAAATGACAGTTTTTCGGGAAACCACTGCTATTTCAACAAATGAGCTTTCACATACCA	1508
OY	1693	GCAACACATATATATTTGTTTCAACGGCTGGGGTTGCCCAAGCTCCCTTTACACCGGTGTGA	1752
Db	1509	GCAACACATATATATTTGTTTCAACGGCTGGGGTTGCCCAAGCTCCCTTTACACCGGTGTGA	1568
OY	1753	TATTTACAGGGAAGGCTGCTGAGTGTGTGCTCGCCCGAAGACCTTCTCTGTGCTTGGGA	1812
Db	1569	TATTTACAGGGAAGGCTGCTGAGTGTGTGCTCGCCCGAAGACCTTCTCTGTGCTTGGGA	1628
OY	1813	TGTTTCTGCAATGTTCTGCTATTTTCCCATGTGAAAGAGACGCAACAAGACCAAGATAT	1872
Db	1629	TGTTTCTGCAATGTTCTGCTATTTTCCCATGTGAAAGAGACGCAACAAGACCAAGATAT	1688
OY	1873	AAGAAATGGAAGCCCACTGACCTCATGTTCAGACTTACCAATGATATCAACCATGGGCA	1932
Db	1689	AAGAAATGGAAGCCCACTGACCTCATGTTCAGACTTACCAATGATATCAACCATGGGCA	1748
OY	1933	CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGATAGTAGACATTTTGGATGCAG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGATAGTAGACATTTTGGATGCAG	1808
OY	1993	TCCGAAGTCCGAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGAAA	2052
Db	1809	TCCGAAGTCCGAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGAAA	1868
OY	2053	AGAAGAGATCAGAAGTGAATGATCATATATCATCAGGACACATCAAGGCTTTCGTACGTAG	2112
Db	1869	AGAAGAGATCAGAAGTGAATGATCATATATCATCAGGACACATCAAGGCTTTCGTACGTAG	1928
OY	2113	TCCTAACACAGAGATTCAAGCAATTAACCTTGCATCGGCTGGAAACATGGGCTTCATACA	2172
Db	1929	TCCTAACACAGAGATTCAAGCAATTAACCTTGCATCGGCTGGAAACATGGGCTTCATACA	1988
OY	2173	AACCTTCTTAAAGGATACCTGGAAATGTATGACACAAGAAATTTGGAAAGAACTTCTTCA	2232
Db	1989	AACCTTCTTAAAGGATACCTGGAAATGTATGACACAAGAAATTTGGAAAGAACTTCTTCA	2048
OY	2233	TAAAGATGATGATGAGATGGCTCTTAAGACCAAAAGAAATGTCTCAATAGACATGACCTTAG	2292
Db	2049	TAAAGATGATGATGAGATGGCTCTTAAGACCAAAAGAAATGTCTCAATAGACATGACCTTAG	2108
OY	2293	CCAGAAGGCTCTGGTACAGAGACTTCATGCAAGCTCATCAACCCCAATCTTCAACAGAT	2352
Db	2109	CCAGAAGGCTCTGGTACAGAGACTTCATGCAAGCTCATCAACCCCAATCTTCAACAGAT	2168

Oy	2353	GGATGAGTTCGTGAAACAAGTTTGGAAAAGGACCGAAAACAAGTCGGCAAAAGCCAGG	2412
Db	2169	GGATGAGTTCGTGAAACAAGTTTGGAAAAGGACCGAAAACAAGTCGGCAAAAGCCAGG	2222
Oy	2413	ACATATCCCAGGGAAACGTATCAATATGAGACCTTACAGAAAATTAAGAAAGTTAGAA	2472
Db	2229	ACATATCCCAGGGAAACGTATCAATATGAGACCTTACAGAAAATTAAGAAAGTTAGAA	2288
Oy	2473	CAGGAGGACCCAGCAATTTTGAGAGGGACCCAGAGAGTCTGAGCTGATTAACCTCTAGA	2533
Db	2289	CAGGAGGACCCAGCAATTTTGAGAGGGACCCAGAGAGTCTGAGCTGATTAACCTCTAGA	2342
Oy	2533	AACCTCAAAACAATGAAACTTGGCTTAGACAATACTGAGAAAAACAATGCAATTAACAT	2592
Db	2349	AACCTCAAAACAATGAAACTTGGCTTAGACAATACTGAGAAAAACAATGCAATTAACAT	2402
Oy	2593	GAACTTTTTCATGGCAATTATGTCGATGTTTACATGGTGGAAATTCAGCTGAGTTTCCA	2652
Db	2409	GAACTTTTTCATGGCAATTATGTCGATGTTTACATGGTGGAAATTCAGCTGAGTTTCCA	2462
Oy	2653	CCAAATTAATAATTAATCCAGATTAACCTTCCATAATAGCTTTTTCCT	2700
Db	2469	CCAAATTAATAATTAATCCAGATTAACCTTCCATAATAGCTTTTTCCT	2516

RESULT 4
 HUMHSEM
 LOCUS HUMHSEM 2530 bp mRNA linear PRI 08-MAY-1995
 DEFINITION Homo sapiens semaphorin-III (Hsema-1) mRNA, complete cds.
 ACCESSION L26081
 VERSION L26081.1 GI:799328
 KEYWORDS semaphorin.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 2530)
 Klotodkin A.L., Matthes, D.J. and Goodman, C.S.
 The semaphorin genes encode a family of transmembrane and secreted
 growth cone guidance molecules
 CELL 75 (7), 1389-1399 (1993)
 JOURNAL MEDLINE 94094332
 PUBMED 8269517
 COMMENT On May 8, 1995 this sequence version replaced gi:436559.
 Original source text: Homo sapiens (human).
 Location/Qualifiers
 1..2530
 FEATURES
 source

```

Gene
/gene="Hsema-III"
1..2530
16..2331
/gene="Hsema-III"
/codon_start=1
/product="gemaphorin-III"
/protein_id="AA065938.1"
/db_xref="GI:436560"
/translation="MGWLTRIVCLFPGVLLTARANYONGKNVPRLLKLSYKEMESNN
VITFNGGLNASSSYHFLIDBERSRLVYAKOHISFIDLVNIKDFOKLWPAVYSYRDESD
CKMAQGLIKCEKCANIKVILKAYNQTHLYACGTGAHPICTYLIELGHHEBDIPIKLEEN
HPENRGSGPYDPKLLTSLILDELGYFTADPENGROFALFRTILGNHNPRTBOCHDE
RWINDPKRTISAHILSESNDPEBDKVFYFERNALDGBEGSACTHARIGOIKNPGEHG
RSIYNKMTTSLIKARLICSVPNGIDIHDEILOVFLNPNPDKVNPVYGGFTTSSNI
EKGSVCMYKSSDURVRLFGYARHAGDGYVWYQGRVYPRCTGSGKGGPDSY
KDLPEDVITTPARSHPMATNPVYFPKANREIYAKTIVNYGFTIIVYDRUNDAEGQIDW
IGTIDVGLAKVASIPKEWTYDLBEVLBEEMVFEFPTAISMEILSKOQVLIIGSLTAC
VAQPLAHACDILYGAKCAECCLARDPYCAMDGSACSARFPFKARTRIRDINNGPLTH
CSDLHDHNDHGHSPERRIIVGVENSTFLECSGQRALVYMOFQRNREEKEIRIVAC
DHIIITDDGLLRSLIQCGKSGNYLCHAVEHGFIQTLTKVLTLEVIDTEHLEILKDDDD

```

GDGSKTKEMSNMTSPQVYTRDEWQLINHPNLTMTDEFCEQWKRDKRQRORPGHT
PENSNNKHJLQENKGRNRRTHEFERARPSV
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN

Query Match 92.6%; Score 2508; DB 9; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGCAGCATGGGCTGTTAACTAGATGTCTGCTTTCTGGGAGTATTACTTACAGC 252
DB 9 CTGCAGCATGGGCTGTTAACTAGATGTCTGCTTTCTGGGAGTATTACTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAGAACATGTGCCAAGGCTGAATTTCTTCAAAAGA 312
DB 69 AAGAGCAAACTATCAGAAATGGGAGAACATGTGCCAAGGCTGAATTTCTTCAAAAGA 128
QY 313 AATGTGGAATTCACATATGTATCACTTTCAATGGCTGGCCACAGCTCCAGTTATCA 372
DB 129 AATGTGGAATTCACATATGTATCACTTTCAATGGCTGGCCACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTGATGAGAACGGAGTAGGCTGATGTGAGAGCAAGATCAATAT 432
DB 189 TACCTTCCTTTGATGAGAACGGAGTAGGCTGATGTGAGAGCAAGATCAATAT 248
QY 433 TTCAATTCAGCTGGTAAATATCAAGATTTTCAAAAGATGTGTGCCAGTATCTTACAC 492
DB 249 TTCAATTCAGCTGGTAAATATCAAGATTTTCAAAAGATGTGTGCCAGTATCTTACAC 308
QY 493 CAGAGGAATGATATGCAAGTGGGCTGGAAGAACATCTCTGAAGATGTCTAATTTGAT 552
DB 309 CAGAGGAATGATATGCAAGTGGGCTGGAAGAACATCTCTGAAGATGTCTAATTTGAT 368
QY 553 CAGAGTACTTAAAGCATATATACAGCTCACTGTAGAGCTGTGGAAGGGGGCTTTTCA 612
DB 369 CAGAGTACTTAAAGCATATATACAGCTCACTGTAGAGCTGTGGAAGGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTATGAAATTTGACATATCTGAGAGCAATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACTATGAAATTTGACATATCTGAGAGCAATATTTTAAAGCTGGA 488
QY 673 GAATCACAATTTGAAAAAGCGCGTGGGAGAGTGCATATGACCTTAAGCTGACAGC 732
DB 489 GAATCACAATTTGAAAAAGCGCGTGGGAGAGTGCATATGACCTTAAGCTGACAGC 548
QY 733 ATCCCTTTAATAGATGGAATTAATCTGGAAGTGCAGAGTGAATTTTAAAGGGCGAGA 792
DB 549 ATCCCTTTAATAGATGGAATTAATCTGGAAGTGCAGAGTGAATTTTAAAGGGCGAGA 608
QY 793 CTTTGTATCTTCCGAACCTTTGGGACCAACCCATTCAGAGCAGAGCATGATTC 852
DB 609 CTTTGTATCTTCCGAACCTTTGGGACCAACCCATTCAGAGCAGAGCATGATTC 668
QY 853 CAGGTGCTCAATGATCCAAAGTTGATAGTCCCACTCATCTCAGAGATGACAAATCC 912
DB 669 CAGGTGCTCAATGATCCAAAGTTGATAGTCCCACTCATCTCAGAGATGACAAATCC 728
QY 913 TGAAGATGACAAAGTATCTTTTCTCCGTGAAATGCAATAGATGGAACACTCTGG 972
DB 729 TGAAGATGACAAAGTATCTTTTCTCCGTGAAATGCAATAGATGGAACACTCTGG 788
QY 973 AAAAGTCTACAGCTAGATATGATGATGCAAGATGATTTGAGGGGACAGAAAG 1032
DB 789 AAAAGTCTACAGCTAGATATGATGATGCAAGATGATCTTTGAGGGGACAGAAAG 848
QY 1033 TCTGTGATTAATGAGCAACATCTCTCAAAAGCTGTCTGATTTGCTCAGTGCAGGTCC 1092
DB 849 TCTGTGATTAATGAGCAACATCTCTCAAAAGCTGTCTGATTTGCTCAGTGCAGGTCC 908
QY 1093 AAAATGCAATGACATCTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1152
DB 909 AAAATGCAATGACATCTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 968

QY 1153 TCTTAAATATCAGCTGTATATGAGATGTTTACGACTTCCAGTAACTTTTCAAGGATC 1212
DB 969 TCTTAAATATCAGCTGTATATGAGATGTTTACGACTTCCAGTAACTTTTCAAGGATC 1028
QY 1213 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
DB 1029 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY 1273 CAGGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTTATCCAGGCTC 1332
DB 1089 CAGGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTTATCCAGGCTC 1148
QY 1333 AGGAATGTGTCAGCAAAACATTTGATGATGATGATGATGATGATGATGATGATGAT 1392
DB 1149 AGGAATGTGTCAGCAAAACATTTGATGATGATGATGATGATGATGATGATGATGAT 1208
QY 1393 TGTATTAACCTTTGCAAGAGTATCCAGCATGATGATGATGATGATGATGATGATGAT 1452
DB 1209 TGTATTAACCTTTGCAAGAGTATCCAGCATGATGATGATGATGATGATGATGATGAT 1268
QY 1453 TCGCCAATAGTATCAAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
DB 1269 TCGCCAATAGTATCAAAACGATGATGATGATGATGATGATGATGATGATGATGAT 1328
QY 1513 AGTGATGCAAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
DB 1329 AGTGATGCAAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
QY 1573 TCTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
DB 1389 TCTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
QY 1633 AGAATGACAGTTTGGGGAACGAGTATGATGATGATGATGATGATGATGATGATGAT 1692
DB 1449 AGAATGACAGTTTGGGGAACGAGTATGATGATGATGATGATGATGATGATGATGAT 1508
QY 1693 GCAACAACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
DB 1509 GCAACAACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
QY 1753 TATTTAGGGGAAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1812
DB 1569 TATTTAGGGGAAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1628
QY 1813 TGGTTGCAAGTCTGCTATTTTCCACGTGAAAGAGCGCAAGAGCAAGATAT 1872
DB 1629 TGGTTGCAAGTCTGCTATTTTCCACGTGAAAGAGCGCAAGAGCAAGATAT 1688
QY 1873 AAGAAATGAGACCACTGACTGACTGCTTACACTTACACATGATATGACATGAGCCA 1932
DB 1689 AAGAAATGAGAGCCACTGACTGACTGCTTACACTTACACATGATATGACATGAGCCA 1748
QY 1933 CAGCCTTGAAGAGATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
DB 1749 CAGCCTTGAAGAGATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
QY 1993 TCCGAATGCGAGAGAGCGCTGCTATTTGGCAATTCAGAGGGAATGAGAGCGAAA 2052
DB 1809 TCCGAATGCGAGAGAGCGCTGCTATTTGGCAATTTCAAGAGGGAATGAGAGCGAAA 1868
QY 2053 AGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
DB 1869 AGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
QY 2113 TCTTAAAGAGATTCAGGCAATTAATCTCTGCTGATGCGGTGGAACATGGGTTCATACA 2172
DB 1929 TCTTAAAGAGATTCAGGCAATTAATCTCTGCTGATGCGGTGGAACATGGGTTCATACA 1988
QY 2173 AACTCTTAAAGTAACTCTGGAAGTATGATGATGATGATGATGATGATGATGATGAT 2232
DB 1989 AACTCTTAAAGTAACTCTGGAAGTATGATGATGATGATGATGATGATGATGATGAT 2048
QY 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292

Db 429 TCCAAATTCGACCTACATTGAAATGGACATCATCTGTGAGCAATATTTTAACTGGA 488
Qy 673 GAACCTCACATTTTGAAGAAAGCGCGTGGGAGAGTCCATATGACCCCTAAGCTCTGACAC 732
Db 489 GAACCTCACATTTTGAAGAAAGCGCGTGGGAGAGTCCATATGACCCCTAAGCTCTGACAC 548
Qy 733 ATCCCTTTTAAATGATGAGAAATTAATCTCTGGAAGTGCAGCTGATTTTATGGGCGAGA 792
Db 549 ATCCCTTTTAAATGATGAGAAATTAATCTCTGGAAGTGCAGCTGATTTTATGGGCGAGA 608
Qy 793 CTTCGTAATCTCCGAACCTTGGGACACACACCCAACTGAGACAGAGCAGACATGATTC 852
Db 609 CTTCGTAATCTCCGAACCTTGGGACACACACCCAACTGAGACAGAGCAGACATGATTC 668
Qy 853 CAGGTGCTCATGATGCCAAAGTTGATAGTCCCACTCATCTGAGAGTGAACAATCC 912
Db 669 CAGGTGCTCATGATGCCAAAGTTGATAGTCCCACTCATCTGAGAGTGAACAATCC 728
Qy 913 TGAAGATGACAAAGTATCTTTTCTCCGTGAAAAATGCAATGATGAGAACTCTGG 972
Db 729 TGAAGATGACAAAGTATCTTTTCTCCGTGAAAAATGCAATGATGAGAACTCTGG 788
Qy 973 AAAACCTACGCTAGATAGTCAAGATATGCAAGAAATGATTTGGAGGCAACAAG 1032
Db 789 AAAACCTACGCTAGATAGTCAAGATATGCAAGAAATGATTTGGAGGCAACAAG 848
Qy 1033 TCTGTGATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTGAGTCCAGGTCC 1092
Db 849 TCTGTGATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTGAGTCCAGGTCC 908
Qy 1093 AAAATGCAATGACACTATTTTATGATGAATGAGAGATGATTCCTAATGAATTTAAAGA 1152
Db 909 AAAATGCAATGACACTATTTTATGATGAATGAGAGATGATTCCTAATGAATTTAAAGA 968
Qy 1153 TCCCTAAAAATCAGTTGATATAGAGTGTTCAGCTTCCAGTACATTTTAAAGGATC 1212
Db 969 TCCCTAAAAATCAGTTGATATAGAGTGTTCAGCTTCCAGTACATTTTAAAGGATC 1028
Qy 1213 AGCCGTGTATGTATAGCATGATGATGAGAGAGGTTCCTTGGTCCATATGCCA 1272
Db 1029 AGCCGTGTATGTATAGCATGATGATGAGAGAGGTTCCTTGGTCCATATGCCA 1088
Qy 1273 CAGGATGAGACCAATCATCAATGAGGTCTTATCAAGAGAGTCCCTATCAACGGCC 1332
Db 1089 CAGGATGAGACCAATCATCAATGAGGTCTTATCAAGAGAGTCCCTATCAACGGCC 1148
Qy 1333 AGGAATCTGTCCACCAAAAATTTGTGTGTTTGAATCTTCAAAAAGACCTTCTGATGA 1392
Db 1149 AGGAATCTGTGTCCACCAAAAATTTGTGTGTTTGAATCTTCAAAAAGACCTTCTGATGA 1208
Qy 1393 TGTATTAACCTTTGCAAGAGTCAATGAGCATGATCAAGTATCAAGTATTCATGACAA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCAATGAGCATGATCAAGTATCAAGTATTCATGACAA 1268
Qy 1453 TCGCCCAATAGTATCAAAAAGATGTAATTTATCAAAATTTGTGATGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAAAGATGTAATTTATCAAAATTTGTGATGACCG 1328
Qy 1513 AGTGAATGCAAGATGAGCAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGAATGCAAGATGAGCAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1388
Qy 1573 TCTTAAGATGTTCAATCTCTAAGAGACTGATGATGTTAAGAGAGGTTCTGCTGGA 1632
Db 1389 TCTTAAGATGTTCAATCTCTAAGAGACTGATGATGTTAAGAGAGGTTCTGCTGGA 1448
Qy 1633 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAAGCAATGAGCTTTCCATAGCA 1692
Db 1449 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAAGCAATGAGCTTTCCATAGCA 1508
Qy 1693 GCAAAATCATATATTTGTTTCAAGCGCTGGGCTTCCCAAGCTCCCTTTTACACCGGTGGA 1752
Db 1509 GCAAAATCATATATTTGTTTCAAGCGCTGGGCTTCCCAAGCTCCCTTTTACACCGGTGGA 1568

Qy 1753 TATTTACGGGAAGAGCTGTGCTGATGTGTGCTGCGCCGAGACCTTACTGTGCTGGGA 1812
Db 1569 TATTTACGGGAAGAGCTGTGCTGATGTGTGCTGCGCCGAGACCTTACTGTGCTGGGA 1628
Qy 1813 TGTGTTCTCATGTTTCCCTATTTTCCACTGCAAAAGAGCGCACAGACAAAGATAT 1872
Db 1629 TGTGTTCTCATGTTTCCCTATTTTCCACTGCAAAAGAGCGCACAGACAAAGATAT 1688
Qy 1873 AAGAAATGAGACCCACGACCTGCTGCTGACCTTACACATGATATACCATGAGGCA 1932
Db 1689 AAGAAATGAGACCCACGACCTGCTGCTGACCTTACACATGATATACCATGAGGCA 1748
Qy 1933 CAGCCCTGAAGAGAAATCATCTATGCTGATGAGATATGATGACATTTTGGATGAG 1992
Db 1749 CAGCCCTGAAGAGAAATCATCTATGCTGATGAGATATGATGACATTTTGGATGAG 1808
Qy 1993 TCCGAATGCGAGAGCGCTGTCTATTTGCAATTCAGAGCGCAAAATGAAAGCGAAA 2052
Db 1809 TCCGAATGCGAGAGCGCTGTCTATTTGCAATTCAGAGCGCAAAATGAAAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGAATGATATCATATGATGAGACAGATCAAGGCTTGTGCTAG 2112
Db 1869 AGAAGAGATCAGAGTGAATGATATCATATGATGAGACAGATCAAGGCTTGTGCTAG 1928
Qy 2113 TCTACAAAGAGATTCAGGCAATTAATCTGACATCGGTGGAACATGGGTTTATCA 2172
Db 1929 TCTACAAAGAGATTCAGGCAATTAATCTGACATCGGTGGAACATGGGTTTATCA 1988
Qy 2173 AACTCTTCTTAAGGTAACCTTGAAGATGATGACACAGAGCATTTGAAAGACTTCTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTTGAAGATGATGACACAGAGCATTTGAAAGACTTCTCA 2048
Qy 2233 TAAAGATGATGAGATGAGTGTCTTAAAGCAAAAGAAATGTCATATGACATGACCTAG 2292
Db 2049 TAAAGATGATGAGATGAGTGTCTTAAAGCAAAAGAAATGTCATATGACATGACCTAG 2108
Qy 2293 CCAGAGGTCTGTGTAACAGAGACTTCAATGACCTCATCAACACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGTGTAACAGAGACTTCAATGACCTCATCAACACCCCAATCTCAACAGAT 2168
Qy 2353 GATGAGTCTGTGTAACAGTGTGAAAAAGGACCGAAAAACAAGTGGCAAGGCGCAGG 2412
Db 2169 GATGAGTCTGTGTAACAGTGTGAAAAAGGACCGAAAAACAAGTGGCAAGGCGCAGG 2228
Qy 2413 ACATACCCAGAGAAACGTATCAAAATGAGAACCTTCAAAAGAAATTAAGAAAGTAA 2472
Db 2229 ACATACCCAGAGAAACGTATCAAAATGAGAACCTTCAAAAGAAATTAAGAAAGTAA 2288
Qy 2473 CAGAGAGACCCAGAAATTTGAGAGGCAACCAAGAGTGTCTGAGCTGATTAACCTTGA 2532
Db 2289 CAGAGAGACCCAGAAATTTGAGAGGCAACCAAGAGTGTCTGAGCTGATTAACCTTGA 2348
Qy 2533 AACCTCAAAACAAGTAAGAACTTGTCTGACAAATTAATGAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAAGTAAGAACTTGTCTGACAAATTAATGAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
Db 2409 GAACCTTTTATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
Qy 2653 CCAATTAATTAATTAATCATGATGATTAATCTTCTAATGAGCTTTTCTT 2700
Db 2469 CCAATTAATTAATTAATCATGATGATTAATCTTCTAATGAGCTTTTCTT 2516

RESULT 6
AR040598 2601 bp DNA linear PAT 29-SEP-1999
LOCUS AR040598
DEFINITION Sequence 53 from patent US 5807826.
ACCESSION AR040598
VERSION AR040598.1 GI:5959961
KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 2601)
 AUTHORS Goodman, C.S., Kolodkin, A.L., Matthes, D., Bentley, D.R. and O'Connor, T.
 TITLE Semaphorin gene family
 JOURNAL Patent: US 5807826-A 53 15-SEP-1998;
 FEATURES
 source 1..2601
 /organism="unknown"
 BASE COUNT 809 a 533 c 593 g 666 t
 ORIGIN
 Query Match 92.6%; Score 2508; DB 6; Length 2601;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 CTGACGATGGGCTGTTAACTAGATGTCTGCTCTTTCTGGGAGTATTAATAACAGC 252
 9 CTGACGATGGGCTGTTAACTAGATGTCTGCTCTTTCTGGGAGTATTAATAACAGC 68
 253 AAGAGCAATCTATCAGAAATGGGAGAAAGATGTCAGGCTGAATTTCTTACAAAGA 312
 69 AAGAGCAATCTATCAGAAATGGGAGAAAGATGTCAGGCTGAATTTCTTACAAAGA 128
 313 AATGTTGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATCA 372
 129 AATGTTGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATCA 188
 373 TACCTTCTTTGATGAGAGAACGAGTGTGATGTTGAGCAAGATCACATATT 432
 189 TACCTTCTTTGATGAGAGAACGAGTGTGATGTTGAGCAAGATCACATATT 248
 433 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATTTCTGCGCAGTATCTTAC 492
 249 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATTTCTGCGCAGTATCTTAC 308
 493 CAGAGAGATGATGCAAGTGGCTGGAAAGACATCTGAAAGATGTCTAATTTCAAT 552
 309 CAGAGAGATGATGCAAGTGGCTGGAAAGACATCTGAAAGATGTCTAATTTCAAT 368
 553 CAGGTAATTAAGCATATATCACTCACTTGAAGCTGTGAGCGGGGCTTTTCA 612
 369 CAGGTAATTAAGCATATATCACTCACTTGAAGCTGTGAGCGGGGCTTTTCA 428
 613 TCCAAATTCGACCTATGAAATGGAGATCATCTGAGGACCAATTTTAAAGCTGA 672
 429 TCCAAATTCGACCTATGAAATGGAGATCATCTGAGGACCAATTTTAAAGCTGA 488
 673 GAATCTACATTTTGAAGACGCGCTGGAGAGATCCATATGACCTTAACTGTCAGAC 732
 489 GAATCTACATTTTGAAGACGCGCTGGAGAGATCCATATGACCTTAACTGTCAGAC 548
 733 ATCCCTTTAATAGATGAGAAATTATATCTGTGAACTGACGCTGATTTATGGGGGAGA 792
 549 ATCCCTTTAATAGATGAGAAATTATATCTGTGAACTGACGCTGATTTATGGGGGAGA 608
 733 CTTCCTATCTCCGAACTCTGGGACACCAATCCAGAGACGACACATGATTC 852
 609 CTTCCTATCTCCGAACTCTGGGACACCAATCCAGAGACGACACATGATTC 668
 853 CAGGTGCTCATGATCCAAAGTCTATAGTCCCACTCATCTCAGAGATGACATATCC 912
 669 CAGGTGCTCATGATCCAAAGTCTATAGTCCCACTCATCTCAGAGATGACATATCC 728
 913 TGAAGATGCAAAAGTATCTTTTCTTCCGTGAAAATGCAATGATGAGAACACTCTGG 972
 729 TGAAGATGCAAAAGTATCTTTTCTTCCGTGAAAATGCAATGATGAGAACACTCTGG 788
 973 AAAAATCTACAGCTAGAAATGTCATATGCAAGATGATGATGAGGAGCAAGAG 1032
 789 AAAAATCTACAGCTAGAAATGTCATATGCAAGATGATGATGAGGAGCAAGAG 848

1033 TCTGTGATTAATATGAGCAACATTCCTCAAGCTGCTGATTTGCTAGTCCAGGTCC 1092
 849 TCTGTGATTAATATGAGCAACATTCCTCAAGCTGCTGATTTGCTAGTCCAGGTCC 908
 1093 AATGCAATGACATCTATTTGATGAATCTGAGAGATGATTTCTTAATGAATTTAAGA 1152
 909 AATGCAATGACATCTATTTGATGAATCTGAGAGATGATTTCTTAATGAATTTAAGA 968
 1153 TCCTAATAATCCAGTTGATATGAGGTTTACAGCTTCCAGTACATTTTCAAGGATC 1212
 969 TCCTAATAATCCAGTTGATATGAGGTTTACAGCTTCCAGTACATTTTCAAGGATC 1028
 1213 ACCCGTGTATGATATGATGATGATGATGAGAGGAGTTCCTTGTCCATATGCCA 1272
 1029 ACCCGTGTATGATATGATGATGATGATGAGAGGAGTTCCTTGTCCATATGCCA 1088
 1273 CAGGATGAGACCAATCTATCAATGGGTCCCTTATCAAGAGATCCCTTATCCAGGCC 1332
 1089 CAGGATGAGACCAATCTATCAATGGGTCCCTTATCAAGAGATCCCTTATCCAGGCC 1148
 1333 AGGAATCTGTCAGCAAAAATTTGGTGTGATCTCAAGAGATCCCTTATCCAGTGA 1392
 1149 AGGAATCTGTCAGCAAAAATTTGGTGTGATCTCAAGAGATCCCTTATCCAGTGA 1208
 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGAACA 1452
 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGAACA 1268
 1453 TCGCCCAATGATGATCAAAAGGATGTAATTAATTAATTTACAAATTTGTGTGACCG 1512
 1269 TCGCCCAATGATGATCAAAAGGATGTAATTAATTAATTTACAAATTTGTGTGACCG 1328
 1513 AGTGATGAGAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1572
 1329 AGTGATGAGAGATGAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1388
 1573 TCTTAAGTATGATCAATTTCTTAAGAGATGATGATGATGATGATGATGATGATGATG 1632
 1389 TCTTAAGTATGATCAATTTCTTAAGAGATGATGATGATGATGATGATGATGATGATG 1448
 1633 AGAATGACATGATTTTGGGAGAACGATGATGATGATGATGATGATGATGATGATGATG 1692
 1449 AGAATGACATGATTTTGGGAGAACGATGATGATGATGATGATGATGATGATGATGATG 1508
 1693 GCAACATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1752
 1509 GCAACATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1568
 1753 TATTTACGGGAAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1812
 1569 TATTTACGGGAAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1628
 1813 TGTGTCGATGATTTTGGGAGAACGATGATGATGATGATGATGATGATGATGATGATG 1872
 1629 TGTGTCGATGATTTTGGGAGAACGATGATGATGATGATGATGATGATGATGATGATG 1688
 1873 AAGAAATGAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1932
 1689 AAGAAATGAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1748
 1933 CAGCCCTGAAGAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1992
 1749 CAGCCCTGAAGAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1808
 1993 TCCGAATGCGAGAGAGGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 2052
 1809 TCCGAATGCGAGAGAGGCTGTATGATGATGATGATGATGATGATGATGATGATGATG 1868
 2053 AGAAGATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2112
 1869 AGAAGATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1928

QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTATCAATCAAGTGTCTATGAAACA 1452
 DB 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTATCAATCAAGTGTCTATGAAACA 1268
 QY 1453 TCGCCCAATGTATGATCAAAACGATGTAAATTTATCAATTTATCAAAATTTGTCTAGAACCG 1512
 DB 1269 TCGCCCAATGTATGATCAAAACGATGTAAATTTATCAATTTATCAAAATTTGTCTAGAACCG 1328
 QY 1513 AGTGAATGCAAGATGAGACAGTATGATTTATGTTATCGAAACAGATTTGGGACCGT 1572
 DB 1329 AGTGAATGCAAGATGAGACAGTATGATTTATGTTATCGAAACAGATTTGGGACCGT 1388
 QY 1573 TCTTAAGTATGTTCAATTCCTTAAGAGACTGTGATGATTTAAGAAAGGTTCTGCTGGA 1632
 DB 1389 TCTTAAGTATGTTCAATTCCTTAAGAGACTGTGATGATTTAAGAAAGGTTCTGCTGGA 1448
 QY 1633 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTCAAGCAATGAGACTTTCCTAAGCA 1692
 DB 1449 AGAAATGACAGTTTTCGGGAACCGACTGCTATTTCAAGCAATGAGACTTTCCTAAGCA 1508
 QY 1693 GCACAACTATATATTTGTTCAACGCGTGGGTTGCGCAGCTCCCTTTACACCGGTGTA 1752
 DB 1509 GCACAACTATATATTTGTTCAACGCGTGGGTTGCGCAGCTCCCTTTACACCGGTGTA 1568
 QY 1753 TATTACGGGAAGCGTGTCTGATGTGTGCTGCGCCGAGACCTTTACTGTCTTGGGA 1812
 DB 1569 TATTACGGGAAGCGTGTCTGATGTGTGCTGCGCCGAGACCTTTACTGTCTTGGGA 1628
 QY 1813 TGGTTCGATGTTCTCGCTATTTTCCACTGCAAGAGCGCAAGACGCAAAATAT 1872
 DB 1629 TGGTTCGATGTTCTCGCTATTTTCCACTGCAAGAGCGCAAGACGCAAAATAT 1688
 QY 1873 AAGAAATGAGACCCCACTGACTCACTGTTCAAGCTTACACCATGATATACCATGAGCA 1932
 DB 1689 AAGAAATGAGACCCCACTGACTCACTGTTCAAGCTTACACCATGATATACCATGAGCA 1748
 QY 1933 CAGCCCTGAAGAGAGATCATCTATGCTGATGAGAAATAGTAGACATTTTGGATGCA 1992
 DB 1749 CAGCCCTGAAGAGAGATCATCTATGCTGATGAGAAATAGTAGACATTTTGGATGCA 1808
 QY 1993 TCGGAAGTGGAGAGAGCGTGTCTATTTGGCAATTCGAGGCGCAATTTGAAGAGGAA 2052
 DB 1809 TCGGAAGTGGAGAGAGCGTGTCTATTTGGCAATTTGCAAGGCGCAATTTGAAGAGGAA 1868
 QY 2053 AGAAGAGTCAAGTGTGATCATATCATCAAGACAGATCAAGGCTTCTGCTAGTAG 2112
 DB 1869 AGAAGAGTCAAGTGTGATCATATCATCAAGACAGATCAAGGCTTCTGCTAGTAG 1928
 QY 2113 TCTACACAGAGAGATTCAGGCAATTAACCTGCTCAATGCGGTGGAACATGGCTCATACA 2172
 DB 1929 TCTACACAGAGAGATTCAGGCAATTAACCTGCTCAATGCGGTGGAACATGGCTCATACA 1988
 QY 2173 AACTCTTCTTAAGTAACTCGGAAGTCAATTAACAAGACATTTGGAAGAACTTTTCA 2232
 DB 1989 AACTCTTCTTAAGTAACTCGGAAGTCAATTAACAAGACATTTGGAAGAACTTTTCA 2048
 QY 2233 TAAAGATGATGATGAGATGCTCTTAAGACCAAGAAATGTCCATATGATGACACTAG 2292
 DB 2049 TAAAGATGATGATGAGATGCTCTTAAGACCAAGAAATGTCCATATGATGACACTAG 2108
 QY 2293 CCAAGAGTCTGTGTACAGAGACTTCATGACGCTCAACACCCCAATCTCAACAGAT 2352
 DB 2109 CCAAGAGTCTGTGTACAGAGACTTCATGACGCTCAACACCCCAATCTCAACAGAT 2168
 QY 2353 GGATGATCTGTGTAAACAAGTTTGAAGAGGACGGAACCAACGTCGCAAGAGGCAAG 2412
 DB 2169 GGATGATCTGTGTAAACAAGTTTGAAGAGGACGGAACCAACGTCGCAAGAGGCAAG 2228
 QY 2413 ACATACCCGAGGAAACAGTAAACAAATGAGACCTTACAAAGAAATTAAGAAAGTGA 2472
 DB 2229 ACATACCCGAGGAAACAGTAAACAAATGAGACCTTACAAAGAAATTAAGAAAGTGA 2288
 QY 2473 CAGGAGGACCAAGAAATTTGAAGAGGACCCAGAGAGTGTGATGATTAACCTCTTGA 2532

DB 2289 CAGGAGGACCCAGATTTGAGAGGCAACCCAGAGTGTCTGAGCTGATTAACCTTGA 2348
 QY 2533 AACCTCAACAGATGAAGAACTTCCCTAGACAAATACTGGAAGAAACAAATGCAATATAC 2592
 DB 2349 AACCTCAACAGATGAAGAACTTCCCTAGACAAATACTGGAAGAAACAAATGCAATATAC 2408
 QY 2593 GAACCTTTTTCATGACATATGTGATGTATTAACAATGTTGGGAATTCAGCTGAGTTCCA 2652
 DB 2409 GAACCTTTTTCATGACATATGTGATGTATTAACAATGTTGGGAATTCAGCTGAGTTCCA 2468
 QY 2653 CCAATTTATTAATTAATTCATGATTAACCTTCTTAATGAGCTTTT 2700
 DB 2469 CCAATTTATTAATTAATTCATGATTAACCTTCTTAATGAGCTTTT 2516
 RESULT 8
 LOCUS 111900 1481 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 1 from Patent US 5416197.
 ACCESSION 111900
 VERSION 111900.1 GI:909343
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 1481)
 AUTHORS Raper J.A. and Luo Y.
 TITLE Antibodies which bind human collapsin
 JOURNAL Patent: US 5416197-A 1 16-MAY-1995;
 FEATURES
 source
 1..1481
 /organism="unknown"
 BASE COUNT 454 a 299 c 345 g 383 t
 ORIGIN
 Query Match 46.3%; Score 1253; DB 6; Length 1481;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1303; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 904 TGACAAATCTGAAAGTGAACAAAGTATCTTTTCTCCGTGAAATGCAATATGATGAGA 963
 DB 58 TGACAAATCTGAAAGTGAACAAAGTATCTTTTCTCCGTGAAATGCAATATGATGAGA 117
 QY 964 AACTCTGGAAGAGTCACTACGCTAGAAATAGTGCAGATATGCAAGATGACTTTGAGG 1023
 DB 118 AACTCTGGAAGAGTCACTACGCTAGAAATAGTGCAGATATGCAAGATGACTTTGAGG 177
 QY 1024 GCACGAAGTGTGTAATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTCAGT 1083
 DB 178 GCACGAAGTGTGTAATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTCAGT 237
 QY 1084 GCGAGTCCAAATGCGATTAACATCTTTTGAATGAGATGAGATGATTTCTTAATGAA 1143
 DB 238 GCGAGTCCAAATGCGATTAACATCTTTTGAATGAGATGAGATGATTTCTTAATGAA 297
 QY 1144 CTTTAAAGATCTTAAATTCAGTGTATGAGAGTGTTAAGCACTTCCAGTAACATTTT 1203
 DB 238 CTTTAAAGATCTTAAATTCAGTGTATATGAGAGTGTTAAGCACTTCCAGTAACATTTT 357
 QY 1204 CAGGAGATCAGCGTGTGTATGATATGATGATGATGATGAGAGGTTCTTGTGTC 1263
 DB 358 CAGGAGATCAGCGTGTGTATGATATGATGATGATGATGAGAGGTTCTTGTGTC 417
 QY 1264 ATATGCCCAACAGGATGAGACCAACTATACATTTGGTCTTATCAAGAAAGTCCCTTA 1323
 DB 418 ATATGCCCAACAGGATGAGACCAACTATACATTTGGTCTTATCAAGAAAGTCCCTTA 477
 QY 1324 TCCAGGCGCAGAACTTGTCCAGCAAAAATTGGTGTGTTTGACTTAACAAAGAGACT 1383
 DB 478 TCCAGGCGCAGAACTTGTCCAGCAAAAATTGGTGTGTTTGACTTAACAAAGAGACT 537
 QY 1384 TCTGTATGATGTTATTAACCTTTGCAAGAGATCATTCAGCATGTACATTCAGTGTTC 1443

FEATURES
* 111040 135938: contig of 24899 bp in length
* 135939 136038: gap of unknown length
* 136039 169408: contig of 33370 bp in length
Location/Qualifiers

			/organism="Homo sapiens"
			/mol_type="genomic DNA"
			/db_xref="taxon:9606"
			/chromosome="7"
			/clone="RP11-506C5"
			. . 1260
			/note="assembly_name:Contig11"
			1361. . 2625
			/note="assembly_name:Contig23"
			2726. . 4145
			/note="assembly_name:Contig34"
			4246. . 5681
			/note="assembly_name:Contig35"
			5782. . 7286
			/note="assembly_name:Contig37"
			7387. . 10841
			/note="assembly_name:Contig40"
			10942. . 15222
			/note="assembly_name:Contig41
			clone_end:SP6
			vector_side:right"
			15323. . 16202
			/note="assembly_name:Contig3"
			16303. . 19851
			/note="assembly_name:Contig42"
			19952. . 24367
			/note="assembly_name:Contig43"
			24468. . 30317
			/note="assembly_name:Contig44"
			30418. . 36826
			/note="assembly_name:Contig45"
			36927. . 43389
			/note="assembly_name:Contig46"
			43490. . 50499
			/note="assembly_name:Contig47"
			50600. . 57156
			/note="assembly_name:Contig48"
			57257. . 65679
			/note="assembly_name:Contig49"
			65780. . 76066
			/note="assembly_name:Contig50"
			76167. . 89540
			/note="assembly_name:Contig51"
			89641. . 110939
			/note="assembly_name:Contig52"
			111040. . 139938
			/note="assembly_name:Contig53"
			136039. . 169408
			/note="assembly_name:Contig54"
			29667 a 29411 g 55592 t 2046 others
			ORIGIN
			Query Match 24.1%; Score 652; DB 2; Length 169408;
			Beet Local Similarity 100.0%; Pred.No. 0;
			Matches 652; Conservative 0; Mismatchas 0; Indels 0; Gaps 0;
OY	2058	AGATCAGGTGGATCATCATCATCAGACGATCAAGCCCTTGCTACGAGTCTAC	2117
Dd	123592	AGATCAGGTGGATCATCATCATCAGACGATCAAGCCCTTGCTACGAGTCTAC	123553
OY	2118	AACAGAAAGATTGAGCAATTACTCTTCGCATGCGGTGAACAATGGGTTTATCAAAATC	2177
Dd	123532	AACAGAAAGATTGAGCAATTACTCTTCGCATGCGGTGAACAATGGGTTTATCAAAATC	123473
OY	2178	TTCCTTAAGSTAAACCTGGAAGTCATTGACACAGAGATTGGAAAGACTTTCATANAAG	2237
Dd	123472	TTCCTTAAGSTAAACCTGGAAGTCATTGACACAGAGATTGGAAAGACTTTCATANAAG	123413

OY	22238	TTGATGATGAGAGATGGCTCTTAAGACCAAGAAATGCTCCAAATGATGACACCTAGCCAGA	2297
Db	123412	ATGATGATGAGAGATGGCTCTTAAGACCAAGAAATGCTCCAAATGATGACACCTAGCCAGA	123353
OY	2298	AGGCTGGTACAGAGACTTTCATGCGACTCATCAACACCCCAATCTCAACACGATGATG	2357
Db	123352	AGGCTGGTACAGAGACTTTCATGCGACTCATCAACACCCCAATCTCAACACGATGATG	123293
OY	2358	AGTTCGTGAAACAAGTTTGGAAAAAGGGACCGAAAACAACGTGGGAAAGGCCAGGACATA	2417
Db	123292	AGTTCGTGAAACAAGTTTGGAAAAAGGGACCGAAAACAACGTGGGAAAGGCCAGGACATA	123233
OY	2418	CCCCGAGGAAACAGTAAACAATGSAAGCACTTACAAGAAATTAAGAAAGGTAGAAACAGA	2477
Db	123232	CCCCGAGGAAACAGTAAACAATGSAAGCACTTACAAGAAATTAAGAAAGGTAGAAACAGA	123173
OY	2478	GGACCCACGAATTGAGAGGGGACCCAGAGAGTGTGAGCTGCATTACCTCTTAGAAACCT	2537
Db	123172	GGACCCACGAATTGAGAGGGGACCCAGAGAGTGTGAGCTGCATTACCTCTTAGAAACCT	123113
OY	2538	CAAAACAATAGAAACTTGCTTGAACAATACTGSAAAAAACAATGCAATATATCATGAACCT	2597
Db	123112	CAAAACAATAGAAACTTGCTTGAACAATACTGSAAAAAACAATGCAATATATCATGAACCT	123053
OY	2598	TTTTTCATGCGATTATGATGATGTTTCAATAGTGGGAAATTCAGCTGAGTTCCACCAAT	2657
Db	123052	TTTTTCATGCGATTATGATGATGTTTCAATAGTGGGAAATTCAGCTGAGTTCCACCAAT	
OY	2658	TATTAATTAATCCATGAGTAACCTTCCATATAGGCTTTTTTCCTAATACC	2709
Db	122992	TATTAATTAATCCATGAGTAACCTTCCATATAGGCTTTTTTCCTAATACC	122941

RESULT 10	
AC006322	179640 bp DNA linear PRI 01-MAR-2002
LOCUS	
DEFINITION	Homo sapiens PAC clone RP5-1060B11 from 7q11.23-q21.1, complete sequence.
ACCESSION	AC006322
VERSION	AC006322.2 GI:4454515
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Sulston,J.E. and Waterston,R.
TITLE	Toward a complete human genome sequence
JOURNAL	Science Res. 8 (11), 1097-1108 (1998)
MEDLINE	99063792
PUBMED	9847074
REFERENCE	2 (bases 1 to 179640)
AUTHORS	Holmes,A. and Gregory,S.
TITLE	The sequence of Hmo sapiens PAC clone RP5-1060B11
JOURNAL	Unpublished (2001)
REFERENCE	3 (bases 1 to 179640)
AUTHORS	Waterston,R.H.
TITLE	Direct Submision
JOURNAL	Submitted (09-JAN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	4 (bases 1 to 179640)
AUTHORS	Waterston,R.
TITLE	Direct Submision
JOURNAL	Submitted (18-MAR-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	5 (bases 1 to 179640)
AUTHORS	Waterston,R.
TITLE	Direct Submision
JOURNAL	Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	6 (bases 1 to 179640)
AUTHORS	Waterston,R.

TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 19, 1999 this sequence version replaced gi:4199987.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@atcson.wustl.edu
----- Summary Statistics
Center project name: H_DU1060B11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see http://www.nhgri.nih.gov/DIR/GRN/CHR7_send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong. VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP5-1060B11; actual end is at base position 179640 of RP5-1060B11.

FEATURES

source

1..179640
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q11.23-q21.1"
/clone="RP5-1060B11"
/clone_lib="RPCI-5"
1..866
/rpt_family="MER2_type"
229..267
/rpt_family="AT_rich"
990..1280
/rpt_family="Alu"
2419..2718
/rpt_family="Alu"
2699..2718
/rpt_family="(A)n"
10069..10158
/rpt_family="Alu"
10297..10328
/rpt_family="AT_rich"
10530..10788
repeat_region

/rpt_family="Alu"
10787..10811
/rpt_family="(CAA)n"
11346..11448
/rpt_family="t2"
11862..11925
/rpt_family="AT_rich"
12013..12048
/rpt_family="(TG)n"
12096..12301
/rpt_family="MIR"
12783..13327
/rpt_family="ERV1"
13537..13583
/note="match to EST BF667677 (NID:gil1941572)"
13955..13979
/rpt_family="AT_rich"
14388..14419
/rpt_family="(A)n"
14810..14832
/rpt_family="AT_rich"
15750..15833
/rpt_family="GA-rich"
16712..16886
/rpt_family="MIR"
17541..17698
/note="match to EST BF667677 (NID:gil1941572)"
18316..18395
/rpt_family="(TATATG)n"
18395..18415
/rpt_family="(T)n"
19588..19885
/rpt_family="Alu"
19826..20402
/note="match to EST AL134679 (NID:g6602866)"
20918..21227
/rpt_family="Alu"
21198..21230
/rpt_family="(A)n"
21224..21825
/note="match to EST BG896972 (NID:gil307213)"
21724..21746
/rpt_family="AT_rich"
21823..21887
/note="match to EST BF667677 (NID:gil1941572)"
25562..25640
/rpt_family="Tc2"
25662..25979
/rpt_family="t1"
25985..26050
/rpt_family="Tc2"
26845..26884
/rpt_family="AT_rich"
26910..27015
/rpt_family="(TA)n"
27573..27873
/rpt_family="Alu"
27931..27965
/rpt_family="AT_rich"
29914..30219
/rpt_family="Alu"
30192..30246
/rpt_family="(A)n"
30765..31032
/note="similar to Homo sapiens EST BF356276 (NID:gil1315350)"
31015..31224
/rpt_family="t1"
31246..31545
/rpt_family="Alu"
31786..31806
/rpt_family="AT_rich"
32009..32310
repeat_region

REFERENCE	MO 63108, USA
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
TITLE	4 (bases 1 to 94027)
JOURNAL	Waterston,R.H. Direct Submission Submitted (08-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 94027)
AUTHORS	Waterston,R.H. Direct Submission Submitted (09-AUG-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	6 (bases 1 to 94027)
AUTHORS	Waterston,R. Direct Submission Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	7 (bases 1 to 94027)
AUTHORS	Waterston,R. Direct Submission Submitted (29-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL	On Aug 9, 2001 this sequence version replaced gi:13992782.
COMMENT	

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:sgreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPEC-11 human BAC library was made from the blood of one male donor, as described by Oseogawa, K., Woon, P. Y., Zhao, B., Prenggen, E., Tatemoto, M., Caranese, J. J. and de Jong, P. J. (1999). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.researchgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
PIBACE3.6.

FEATURES
SOURCE

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
```

variation	repeat_region	/clone="RP11-543D8" /clone_lib="RCCT-11" 895 /allele="C" /allele="A" /db_xref="dbSNP:2527038" 1735..1810 /rpe_family="CT-rich" 1811..2295 /rpe_family="Alu" 2342 /allele="G" /allele="A" /db_xref="dbSNP:2691700" 2346..2377 /rpe_family="AT-rich" 2404..2431 /rpe_family="AT-rich" 3092..3122 /rpe_family="AT-rich" 4054 /allele="C" /allele="T" /db_xref="dbSNP:2040877" 4517 /allele="C" /allele="T" /db_xref="dbSNP:2040876" 4611..4905 /rpe_family="Alu" 5376 /allele="G" /allele="A" /db_xref="dbSNP:2691699" 5571..5694 /rpe_family="L2" 5718..5864 /rpe_family="MIR" 5967 /allele="C" /allele="T" /db_xref="dbSNP:2527037" 6318..6370 /rpe_family="ERV1" 6381..6417 /rpe_family="TA)n" 6480..6756 /rpe_family="TA)n" 6788..6885 /rpe_family="MIR" 6912 /allele="A" /allele="T" /db_xref="dbSNP:2527036" 6976 /allele="T" /allele="C" /db_xref="dbSNP:2527035" 7485..7510 /rpe_family="AT-rich" 7578 /allele="G" /allele="A" /db_xref="dbSNP:1524880" 7746 /allele="T" /allele="C" /db_xref="dbSNP:1524881" 8912 /allele="A" /allele="G" /db_xref="dbSNP:2691698" 9493
-----------	---------------	---

```

repeat_region      /allele="G"
                   /allele="C"
                   /db_xref="dbSNP:2527034"
                   9525..9862 "L1"
variation           /rpt_family="L1"
                   /allele="C"
                   /allele="A"
                   /db_xref="dbSNP:2189847"
repeat_region      /rpt_family="AT_rich"
                   10203..10227
repeat_region      /rpt_family="A-rich"
                   11126..11174
repeat_region      /rpt_family="A-rich"
                   11550..11855
repeat_region      /rpt_family="Alu"
                   12441..12462
repeat_region      /rpt_family="AT_rich"
                   12475..13649
repeat_region      /rpt_family="L1"
                   13919
variation           /allele="C"
                   /allele="T"
                   /db_xref="dbSNP:2040875"
repeat_region      /rpt_family="A-rich"
                   14069..14124
repeat_region      /rpt_family="A-rich"
                   14391..14438
variation           /rpt_family="MER1_type"
                   14460
                   /allele="G"
                   /allele="T"
                   /db_xref="dbSNP:2107738"

Query Match      8.1% Score 220; DB 9; Length 94027;
Best Local Similarity 100.0%; Pred. No. 4,2e-113;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1340 TGTCCAGCAAAACATTTGGTGTGTTGACTCTACAAAGACCTTCTGATGATGTATA 1399
DB 87288 TGTCCAGCAAAACATTTGGTGTGTTGACTCTACAAAGACCTTCTGATGATGTATA 87347

QY 1400 ACCTTTGCAAGAGTATCCAGCCATGTCATCCAGTGTTCCTTCTTGAACAATGCCCA 1459
DB 87348 ACCTTTGCAAGAGTATCCAGCCATGTCATCCAGTGTTCCTTCTTGAACAATGCCCA 87407

QY 1460 ATAGAGTATCAAAACGAGTGAATTTCAATTATTAACAATGTCGAGCGAGTGAT 1519
DB 87408 ATAGAGTATCAAAACGAGTGAATTTCAATTATTAACAATGTCGAGCGAGTGAT 87467

QY 1520 GCAGAGATGACAGATGATGTTATGTTATCGAAGAG 1559
DB 87468 GCAGAGATGACAGATGATGTTATGTTATCGAAGAG 87507

RESULT 13
AC004451/c 108642 bp DNA linear PRI 04-JUN-2002
LOCUS Homo sapiens PAC clone RP4-789N1 from 7q21, complete sequence.
DEFINITION AC004451
ACCESSION AC004451
VERSION AC004451.2 GI:21322198
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 108642)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 108642)
AUTHORS Woessner,J., Minx,P., Hands,K. and Stromwater,C.
TITLE The sequence of Homo sapiens PAC clone RP4-789N1

```

```

JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 108642)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 108642)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 108642)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 4, 2002 this sequence version replaced gi:2979604.
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
Center project name: H_DJ0789N01
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPI-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP4-649P17; the clone sequenced to the right is RP4-800B9, 200 bp overlap. Actual start of this clone is at base position 78831 of RP4-649P17; actual end is at base position 108516 of RP4-789N1.

FEATURES

source

```

1..108642
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21"
/clone="RP4-789N1"
/clone_lib="RPI-4"

```



```
repeat_region 1366..1394 /rpt_family="AT_rich"
repeat_region 3626..3692 /rpt_family="(GA)n"
repeat_region 4043..4087 /rpt_family="(CA)n"
repeat_region 4975..5033 /rpt_family="(GA)n"
repeat_region 5063..5123 /rpt_family="purine-rich"
repeat_region 5481..5519 /rpt_family="AT_rich"
repeat_region 6368..6676 /rpt_family="Alu"
repeat_region 6820..7138 /rpt_family="Alu"
repeat_region 10797..10996 /rpt_family="Alu"
repeat_region 11925..12010 /rpt_family="L2"
repeat_region 12140..12164 /rpt_family="(TA)n"
repeat_region 12165..12467 /rpt_family="Alu"
repeat_region 12793..12823 /rpt_family="AT_rich"
repeat_region 14958..15258 /rpt_family="Alu"
repeat_region 15334..15386 /rpt_family="Alu"
repeat_region 16602..16720 /rpt_family="Alu"
repeat_region 16808..17098 /rpt_family="(TA)n"
repeat_region 17939..17978 /rpt_family="Alu"
repeat_region 18063..18146 /rpt_family="(CA)n"
repeat_region 18419..18453 /rpt_family="MIR"
repeat_region 18802..19345 /rpt_family="AT_rich"
repeat_region 19352..19382 /rpt_family="L2"
repeat_region 19367..19692 /rpt_family="POLY_A"
repeat_region 19367..19692 /note="match to EST AA731945 (NID:g2753896) nw64g09.s1"
misc_feature 19367..19692 /note="similar to EST AA644250 (NID:g2569468) af63b06.s1"
misc_feature 19368..19727 /note="similar to EST AA79841 (NID:g2839172) af45g01.s1"
misc_feature 19378..19791 /note="similar to EST A1122887 (NID:g3538653) qb01f04.s1"
misc_feature 19577..19791 /note="match to EST AA320720 (NID:g1973048)"
misc_feature 19611..19791 /note="similar to EST AA171850 (NID:g1751104) z098f08.s1"
repeat_region 19966..20271 /rpt_family="Alu"
repeat_region 21538..21564 /rpt_family="AT_rich"
repeat_region 21605..21913 /rpt_family="Alu"
repeat_region 21924..22322 /rpt_family="Alu"
repeat_region 22714..22884 /rpt_family="MIR"
repeat_region 23335..23636 /rpt_family="Alu"
repeat_region 23637..23807 /rpt_family="Alu"
repeat_region 23813..23835 /rpt_family="(TAA)n"
repeat_region 23845..24126
```

```
repeat_region /rpt_family="Alu"
repeat_region 24156..24398 /rpt_family="L2"
repeat_region 24717..25041 /rpt_family="Alu"
repeat_region 26317..26414 /rpt_family="Alu"
repeat_region 26433..26627 /rpt_family="Alu"
repeat_region 26703..26729 /rpt_family="AT_rich"
repeat_region 26894..26947 /rpt_family="AT_rich"
repeat_region 27824..27859 /rpt_family="AT_rich"
repeat_region 28102..28368 /rpt_family="MER2_type"
repeat_region 28341..28500 /rpt_family="MER2_type"
repeat_region 28500..30640 /rpt_family="MER2_type"
repeat_region 31284..31318 /rpt_family="AT_rich"
repeat_region 31447..31584 /rpt_family="MER1_type"
repeat_region 31928..32008 /rpt_family="(CAT)n"
repeat_region 35757..35850 /rpt_family="(TA)n"
repeat_region 36529..36559 /rpt_family="(CA)n"
repeat_region 37257..37640 /rpt_family="Retroviral"
repeat_region 38296..38328 /rpt_family="AT_rich"
repeat_region 38501..39044
```

Query Match 7.2%; Score 194; DB 9; Length 108642;
Best Local Similarity 100.0%; Pred. No. 2,9e-98;
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 118 TTTCAGATTGTTGAACTTCTCTGCGCAGCAATACAGAGAGACATAAGCAGCAA 177
Db 3305 TTTCAGATTGTTGAACTTCTCTGCGCAGCAATACAGAGAGACATAAGCAGCAA 3246
Qy 178 AGGACCTACAGCGCTGCGACGATGGCGCTGTTAAGATTTCTCTCTTTCTGGG 237
Db 3245 AGGACCTACAGCGCTGCGACGATGGCGCTGTTAAGATTTCTCTCTTTCTGGG 3186
Qy 238 AGTATTACTTACAGCAGAGCAAACTATCAGAAATGGAGAGACATGTGCCAAGCTGAA 297
Db 3185 AGTATTACTTACAGCAGAGCAAACTATCAGAAATGGAGAGACATGTGCCAAGCTGAA 3126
Qy 298 ATTATCTTACAAAG 311
Db 3125 ATTATCTTACAAAG 3112
```

RESULT 14
AC004848/c 135214 bp DNA linear PRI 26-APR-2003
LOCUS Homo sapiens PAC clone RP4-649P17 from 7, complete sequence.
DEFINITION AC004848
ACCESSION AC004848
VERSION AC004848.1 GI:3900858
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 135214)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 135214)
 AUTHORS Duckels,G., Rohlfing,T. and Tin-Wollam,A.M.
 TITLE The sequence of Homo sapiens PAC clone RP4-649P17
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 135214)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 5 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 6 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Nov 21, 1998 this sequence version replaced gi:3213149.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_DJ0649P17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at http://www.chori.org using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
 The clone may be obtained either from Genome Systems, Inc. (http://www.genomesystems.com) or Research Genetics, Inc. (http://www.resgen.com); or from Pieter de Jong.
 VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP4-789N1, 200 bp overlap.
 Actual start of this clone is at base position 1 of RP4-649P17

FEATURES actual end is at 135018 of DJ0649P17.
 source Location/Qualifiers
 1..135214
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="RP4-649P17"
 /clone_1db="RPCI-4"
 261..298
 /rpt_family="CA)n"
 483
 /rpt_family="CA)n"
 repeat_region
 598..691
 /db_xref="dbSNP:2527040"
 /rpt_family="MIR"
 1021
 /allele="T"
 /allele="C"
 /db_xref="dbSNP:2527030"
 1136..1402
 /rpt_family="Alu"
 1403..1443
 /rpt_family="(TAA)n"
 1957
 /allele="C"
 /allele="T"
 /db_xref="dbSNP:2527032"
 3187..3240
 /rpt_family="AT_rich"
 3415..3465
 /rpt_family="A-rich"
 3525
 /allele="G"
 /allele="A"
 /db_xref="dbSNP:2189848"
 3674
 /allele="T"
 /allele="A"
 /db_xref="dbSNP:1357813"
 3941
 /allele="C"
 /allele="T"
 /db_xref="dbSNP:2301890"
 3949
 /allele="C"
 /allele="T"
 /db_xref="dbSNP:2527039"
 complement(<4116..>78602)
 /gene="SEMA3A"
 complement(join(4116..4209,54121..54240,72774..72836,78445..78602))
 /gene="SEMA3A"
 complement(join(4116..4209,54121..54240,72774..72836,78445..>78602))
 /gene="SEMA3A"
 /note="Homo sapiens sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A (19), short basic domain, secreted, (semaphorin) 3A (SEMA3A), mRNA, H_DJ0649P17.2
 Continues from H_DJ0789N01.1
 Continues as H_NH0543D08.1"
 /codon_start=3
 /product="unknown"
 /protein_id="AAC78622.1"
 /db_xref="GI:3900859"
 /translation="MLSSNVITFNGLAASSYHTFLDEERSRLVYGAKDIFSLVNIKDFQIVPVSVYTRDECKAKGDIKCANFIVKLVKAYNOTHTVYACGTGAHPPICTVIEIGHHPEDNIFKLENSHFENGSGSPYDPKLTPLASLLI"
 4457
 /allele="A"
 /allele="G"
 variation

```

variation      /db xref="dbSNP:1157309"
                5036
                /allele="A"
                /allele="C"
                /db xref="dbSNP:2691702"
variation      5078
                /allele="A"
                /allele="G"
                /db xref="dbSNP:2527041"
                5723. .5751
repeat_region  /rpt_family="AT_rich"
                5906
variation      /allele="C"
                /allele="A"
                /db xref="dbSNP:2189849"
                5953. .5973
repeat_region  /rpt_family="AT_rich"
                5978
variation      /allele="A"
                /allele="G"
                /db xref="dbSNP:2189850"
                6078. .6504
misc_feature   /note="CpG_island (%GC=59.3, o/e=0.89, #CpGs=33)"
                6095
variation      /allele="A"
                /allele="G"
                /db xref="dbSNP:2691701"
                6127
variation      /allele="G"
                /allele="C"
                /db xref="dbSNP:2527042"
                12512
variation      /allele="T"
                /allele="C"
                /db xref="dbSNP:2527031"
                12851. .12970
repeat_region /rpt_family="MIR"
                14674
variation      /allele="C"
                /allele="T"
                /db xref="dbSNP:2888220"
                /db_xref="dbSNP:2888220"

Query Match    5.9%; Score 160; DB 9; Length 135214;
Best Local Similarity 100.0%; Pred. No. 7.4e-79;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGATCCAAACATGTGATCATCTTCAATGGCTTGGCCAAACAGCTCCAGTTA 369
    |||||
Db 78604 AGAATGTTGGATCCAAACATGTGATCATCTTCAATGGCTTGGCCAAACAGCTCCAGTTA 78545

QY 370 TCATACCTTCCTTTGGATGAGGAGAGAGTAGAGCTGATGTTGGAGGAAAGATCACAT 429
    |||||
Db 78544 TCATACCTTCCTTTGGATGAGGAGAGAGTAGAGCTGATGTTGGAGGAAAGATCACAT 78485

QY 430 ATTTTCATTGCACTGTTAATATCAAGATTTTCAAAAG 469
    |||||
Db 78484 ATTTTCATTGCACTGTTAATATCAAGATTTTCAAAAG 78445

RESULT 15
AC073110      169408 bp   DNA      linear   HTG 03-FEB-2001
LOCUS         Homo sapiens chromosome 7 clone RP11-506C5, WORKING DRAFT SEQUENCE,
DEFINITION    21 unordered pieces.
ACCESSION     AC073110
VERSION        AC073110.5 GI:12658189
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE     1 (bases 1 to 169408)
AUTHORS       Waterston,R.H.

```

```

TITLE          The sequence of Homo sapiens clone
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 169408)
AUTHORS        Waterston,R.H.
TITLE          Direct Submission
JOURNAL        Submitted (08-JUN-2000) Genome Sequencing Center, Washington
                University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                MO 63108, USA
COMMENT        On Feb 3, 2001 this sequence version replaced gi:11192243.

-----Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
-----Project Information -----
Center project name: H.NH0506C05
-----Summary Statistics -----
Sequencing vector: M13; 84%
Sequencing strategy: Dye-terminator Big Dye; 5% of reads
Chemistry: Dye-terminator Big Dye; 5% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 157223 bases at least Q40
Consensus quality: 160650 bases at least Q20
Consensus quality: 162652 bases at least Q20
Insert size: 155000; agarose-fp
Quality coverage: 4.21 in Q20 bases; agarose-fp
Quality coverage: 4.09 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      1260: contig of 1260 bp in length
*      1261      1360: gap of unknown length
*      1361      2625: contig of 1265 bp in length
*      2626      2725: gap of unknown length
*      2726      4145: contig of 1420 bp in length
*      4146      4245: gap of unknown length
*      4246      5681: contig of 1436 bp in length
*      5682      5781: gap of unknown length
*      5782      7286: contig of 1505 bp in length
*      7287      7387      10841: contig of 3455 bp in length
*      10842      10941: gap of unknown length
*      10942      15222: contig of 4281 bp in length
*      15223      15322: gap of unknown length
*      15323      16202: contig of 880 bp in length
*      16203      16302: gap of unknown length
*      16303      19851: contig of 3549 bp in length
*      19852      19951: gap of unknown length
*      19952      24367: contig of 4416 bp in length
*      24368      24467: gap of unknown length
*      24468      30317: contig of 5850 bp in length
*      30318      30417: gap of unknown length
*      30418      36826: contig of 6409 bp in length
*      36827      36926: gap of unknown length
*      36927      43388: contig of 6463 bp in length
*      43389      43489: gap of unknown length
*      43490      50499: contig of 7010 bp in length
*      50500      50599: gap of unknown length
*      50600      57156: contig of 6557 bp in length
*      57157      57256: gap of unknown length
*      57257      65679: contig of 8423 bp in length
*      65680      65779: gap of unknown length
*      65780      76066: contig of 10287 bp in length
*      76067      76167      89540: gap of unknown length
*      89540      13374: contig of 13374 bp in length

```



```

* 3359 3458: gap of 100 bp
* 3459 4520: contig of 1062 bp in length
* 4521 4620: gap of 100 bp
* 4621 5725: contig of 1105 bp in length
* 5726 5825: gap of 100 bp
* 5826 6895: contig of 1070 bp in length
* 6896 6995: gap of 100 bp
* 6996 8043: contig of 1048 bp in length
* 8044 8143: gap of 100 bp
* 8144 9212: contig of 1069 bp in length
* 9213 9312: gap of 100 bp
* 9313 10439: contig of 1127 bp in length
* 10440 10539: gap of 100 bp
* 10540 11612: contig of 1073 bp in length
* 11613 11712: gap of 100 bp
* 11713 12750: contig of 1038 bp in length
* 12751 12850: gap of 100 bp
* 12851 13886: contig of 1036 bp in length
* 13887 13986: gap of 100 bp
* 13987 15027: contig of 1041 bp in length
* 15028 15127: gap of 100 bp
* 15128 16197: contig of 1070 bp in length
* 16198 16297: gap of 100 bp
* 16298 17398: contig of 1101 bp in length
* 17399 17498: gap of 100 bp
* 17499 18610: contig of 1112 bp in length
* 18611 18710: gap of 100 bp
* 18711 19754: contig of 1044 bp in length
* 19755 19855: gap of 100 bp
* 19855 20944: contig of 1090 bp in length
* 20945 21044: gap of 100 bp
* 21045 22045: contig of 1000 bp in length
* 22046 22144: gap of 100 bp
* 22145 23221: contig of 1077 bp in length
* 23222 23321: gap of 100 bp
* 23322 24304: contig of 1083 bp in length
* 24305 25433: contig of 1139 bp in length
* 25434 25743: gap of 100 bp
* 25744 26860: contig of 1117 bp in length
* 26861 26960: gap of 100 bp
* 26961 28028: contig of 1068 bp in length
* 28029 28128: gap of 100 bp
* 28129 29216: contig of 1088 bp in length
* 29217 29316: gap of 100 bp
* 29317 30393: contig of 1077 bp in length
* 30394 30493: gap of 100 bp
* 30494 31584: contig of 1091 bp in length
* 31585 31685: gap of 100 bp
* 31685 32761: contig of 1077 bp in length
* 32762 32861: gap of 100 bp
* 32862 33951: contig of 1090 bp in length
* 33952 34051: gap of 100 bp
* 34052 35075: contig of 1024 bp in length
* 35076 35175: gap of 100 bp
* 35176 36272: contig of 1097 bp in length
* 36273 36372: gap of 100 bp
* 36373 37441: contig of 1069 bp in length
* 37442 37541: gap of 100 bp
* 37542 38600: contig of 1059 bp in length
* 38601 38700: gap of 100 bp
* 38701 39781: contig of 1081 bp in length
* 39782 40979: contig of 1098 bp in length
* 40980 41079: gap of 100 bp
* 41080 42096: contig of 1017 bp in length
* 42097 42196: gap of 100 bp
* 42197 43295: contig of 1099 bp in length
* 43296 43395: gap of 100 bp
* 43396 44461: contig of 1066 bp in length
* 44462 44561: gap of 100 bp
* 44562 45625: contig of 1064 bp in length
* 45626 45725: gap of 100 bp

```

```

* 45726 46830: contig of 1105 bp in length
* 46831 46930: gap of 100 bp
* 46931 48033: contig of 1103 bp in length
* 48034 48133: gap of 100 bp
* 48134 49181: contig of 1048 bp in length
* 49182 49281: gap of 100 bp
* 49282 50381: contig of 1100 bp in length
* 50382 50481: gap of 100 bp
* 50482 51600: contig of 1119 bp in length
* 51601 51700: gap of 100 bp
* 51701 52817: contig of 1117 bp in length
* 52818 52917: gap of 100 bp
* 52918 53967: contig of 1050 bp in length
* 53968 54067: gap of 100 bp
* 54068 55103: contig of 1036 bp in length
* 55104 55203: gap of 100 bp
* 55204 56264: contig of 1061 bp in length
* 56265 56364: gap of 100 bp
* 56365 57469: contig of 1105 bp in length
* 57470 57569: gap of 100 bp
* 57570 58636: contig of 1067 bp in length
* 58637 58736: gap of 100 bp
* 58737 59797: contig of 1061 bp in length
* 59798 59897: gap of 100 bp
* 59898 60979: contig of 1082 bp in length
* 60980 61079: gap of 100 bp
* 61080 62036: contig of 956 bp in length
* 62037 62135: gap of 100 bp
* 62136 63164: contig of 1029 bp in length
* 63165 63264: gap of 100 bp
* 63265 64305: contig of 1041 bp in length
* 64306 64405: gap of 100 bp
* 64406 65427: contig of 1022 bp in length
* 65428 65527: gap of 100 bp
* 65528 66574: contig of 1047 bp in length
* 66575 66674: gap of 100 bp
* 66675 67683: contig of 1009 bp in length
* 67684 67783: gap of 100 bp
* 67784 68876: contig of 1093 bp in length
* 68877 68976: gap of 100 bp
* 68977 69998: contig of 1022 bp in length
* 69999 70098: gap of 100 bp
* 70099 71149: contig of 1051 bp in length
* 71150 71249: gap of 100 bp
* 71250 72339: contig of 1090 bp in length
* 72340 72439: gap of 100 bp
* 72440 73511: contig of 1072 bp in length
* 73512 73611: gap of 100 bp
* 73612 74726: contig of 1115 bp in length
* 74727 74826: gap of 100 bp
* 74827 75864: contig of 1038 bp in length
* 75865 75964: gap of 100 bp
* 75965 77015: contig of 1051 bp in length
* 77016 77115: gap of 100 bp
* 77116 78183: contig of 1068 bp in length
* 78184 78283: gap of 100 bp
* 78284 79298: contig of 1015 bp in length
* 79299 79398: gap of 100 bp
* 79399 80451: contig of 1053 bp in length
* 80452 80551: gap of 100 bp

```

Query Match 4.3%; Score 117; DB 2; Length 81667;
 Best Local Similarity 100.0%; Pred. No. 2,7e-54; Indels 0; Gaps 0;
 Matches 117; Conservative 0; Mismatches 0;

```

QY 1008 AGAATGACTTTGAGGGCAGAGAGTGTGTGATTAATGACACATTCTCAAGCTC 1067
DB 20163 AGAATGACTTTGAGGGCAGAGAGTGTGTGATTAATGACACATTCTCAAGCTC 20104
QY 1068 GTCTGATTTCTAGTCCAGTCCCAATGCACTTGTGACACTCTTTTGATGAAGTGC 1124
DB 20103 GTCTGATTTCTAGTCCAGTCCCAATGCACTTGTGACACTCTTTTGATGAAGTGC 20047

```

```

RESULT 17
AX207159      2137 bp      DNA      linear      PAT 30-AUG-2001
LOCUS
DEFINITION    Sequence 6 from Patent WO0155455.
VERSION       AX207159.1  GI:15394947
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS      Jin, S.
TITLE        Resistance sequences and uses thereof.
JOURNAL      Patent: WO 0155455-A 6 02-AUG-2001;
              Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)
FEATURES
source
1. .2137
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
BASE COUNT    658 a      454 c      428 g      597 t
ORIGIN
Query Match   2.6%; Score 71; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 6,1e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGTTTAAACAAGCTTAAATCGATGCGACGCGGCGTGC 60
    |||||||
DB 17 AATCTTTATTTATGATGTTTAAACAAGCTTAAATCGATGCGACGCGGCGTGC 76
    |||||||
QY 61 ACCCAAGCGTC 71
    |||||||
DB 77 ACCCAAGCGTC 87
    |||||||

RESULT 18
RNSIIIC1      2319 bp      mRNA      linear      ROD 29-JAN-1996
LOCUS
DEFINITION    R.norvegicus mRNA for semaphorin III/collapsin-1.
VERSION       X95286.1  GI:1171618
KEYWORDS     collapsin-1; semaphorin III.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.

REFERENCE
AUTHORS      Giger, R.I.
JOURNAL      Unpublished
REFERENCE     2 (bases 1 to 2319)
AUTHORS      Giger, R.I.
TITLE        Direct Submision
JOURNAL      Submitted (22-JAN-1996) R.I. Giger, Netherlands Inst. for Brain
              Research, Meibergdreef 33, 1105 AZ, Amsterdam, NETHERLANDS
FEATURES
source
1. .2319
   /organism="Rattus norvegicus"
   /mol_type="mRNA"
   /strain="Wistar"
   /db_xref="taxon:10116"
   /tissue_type="brain"
   /dev_stage="adult"
1. .2319
   /codon_start=1
   /product="semaphorin III/collapsin-1"
   /protein_id="CAA64607.1"
   /db_xref="GI:1171618"
   /db_xref="SPTREMBL:O63548"
   /translation="MGWFTGIACLFWGIILTARAVYANGKNVPRUKSYKMEESNN"

CDS

```

```

VITFNGLANSSVHTFLDEERSRLVYGAKDHIESFNLWNKDFQKIVMPVSYTRDE
CKMAKDKILKECANIKVLKAYNOTHLVACGAFHPICTVLEVGHEDNIFKQDS
HFNENGRKSPYDPKLTASLIDGLVYGTADFGPRFALFRTLGHEDNIFKQDS
RWLNDPRISAHILIDSDNPRDDKTYFERNALIDGSGKATARIGOLICNDGSH
KSLVKNWTFPLKALICVPGPNGLIDHFDLQVDFLMSKDPKRPITYGFTTSSNI
FKGSAVCMYSMSDVARVFLGPAHRDGNVQGVVPRPQETCSKTFGGSDST
KDLPDVITFAVSHPMAMPVPEFINNRPIMKITDVNVOFTQIVDRAVAEDGQVDMF
IGTDVGLTVKSVKPEWMDLEEVILEEMTVFREPPTISAMELSTKQOQYIGSTAG
VAOLPILRHCDIYGKACAECCCLADPYCAMDSSCSRYPPTAKRRTRRODIRNGDLTH
CSDLOHNDHNDHSHLEERIIYGVENSSTFLCSPSOPALVYMORRNRNDRKERIRV
GDHITRTQGLILSLSLQKDSGNVICHAVHGFNOTLKTLEVIDTTHLESLHKDD
DGDGSKTKEMSSSMTPSQVWYRDMQINHPNLTMDPECEQWKRKORRORPQH
SQGSSNKKWKMHWQESKGRNRTTHERPRSV"

BASE COUNT    704 a      491 c      548 g      576 t
ORIGIN
Query Match   2.5%; Score 68; DB 10; Length 2319;
Best Local Similarity 100.0%; Pred. No. 3,1e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TACAAGAAATGTGGAATCCACAATGATGATTCATTGCGTGGCCACAGCTCC 364
    |||||||
DB 106 TACAAGAAATGTGGAATCCACAATGATGATTCATTGCGTGGCCACAGCTCC 165
    |||||||
QY 365 AGTTATCA 372
    |||||||
DB 166 AGTTATCA 173
    |||||||

RESULT 19
BD026302      229 bp      DNA      linear      PAT 27-AUG-2002
LOCUS
DEFINITION    Sequence tag and encoded human protein.
ACCESSION    BD026302
VERSION      BD026302.1  GI:22567525
KEYWORDS     JP 2001269182-A/2548.
SOURCE       Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS      Edwards, J.B.D.M., Duclair, E. and Jordan, J.V.
JOURNAL      Sequence tag and encoded human protein
              Patent: JP 2001269182-A 2548 02-OCT-2001;
              GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/2548
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC
FH Key Location/Qualifiers
FT CDS 29..229.
1. .229
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
BASE COUNT    63 a      47 c      53 g      62 t      4 others
ORIGIN
Query Match   2.3%; Score 63; DB 6; Length 229;
Best Local Similarity 100.0%; Pred. No. 2,5e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 GCAAGGACCTACAGCGTTCGACAGATGGCGTGGTAACTAGATGCTGCTTTCT 233
    |||||||

```

Db 123 GCAAGGACCTACAGCGCTGCAGCAGCGCTGTTAACTAGATTGTCGCTTTCT 182
 GY 234 GGG 236
 |||
 Db 183 GGG 185

RESULT 20
 AC109862.2/c
 WPCOMMENT

Sequence split into 4 fragments LOCUS AC109862 Accession AC109862

Fragment Name	Begin	End
AC109862_0	1	110000
AC109862_1	100001	210000
AC109862_2	200001	310000
AC109862_3	300001	381240

Continuation (3 of 4) of AC109862 from base 200001 (AC109862 Rattus norvegicus clone CH2)

Query Match 2.3%; Score 63; DB 2; Length 110000;
 Best Local Similarity 100.0%; Pred. No. 1.8e-23;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GY 310 AGAATGTTGAAATCCAAATGATGATCATTTTCATGCTTGGCCACAGCTCCAGTTA 369
 |||
 Db 2015 AGAATGTTGAAATCCAAATGATGATCATTTTCATGCTTGGCCACAGCTCCAGTTA 20056
 |||
 GY 370 TCA 372
 |||
 Db 20055 TCA 20053

RESULT 21
 AC091270
 LOCUS AC091270 81667 bp DNA linear HTG 12-APR-2001
 DEFINITION Homo sapiens chromosome 11 clone RP11-115M2 map 11, LOW-PASS
 SEQUENCE SAMPLING.

AC091270
 AC091270.1 GI:13605997
 HTG: HTGS PHASE0.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 81667)
 Birten,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-115M2
 Unpublished
 2 (bases 1 to 81667)

REFERENCE AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barta,N., Bascien,T., Boguslavsky,L., Bouckgeater,B., Brown,A.,
 Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
 Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
 Diaz,T.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Girard-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
 Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRoque,K.,
 Llamas,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
 Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
 McGowan,P., McKernan,K., McPheters,R., Melidim,J., Menus,L.,
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norby,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
 Reta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnapack,R., Seaman,S.,
 Severi,P., Sougniez,C., Spencer,B., Stange-Thomann,N.,
 Stefanovic,N., Straus,N., Subramanian,A., Talama,J., Testaye,S.,
 Theodore,J., Travers,M., Travers,N., Triggillo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT
 Submitted (12-APR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:

Smc, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIGR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L13092
 Center clone name: 115_M_2

NOTE: This record contains 70 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1	1061:	contig of 1061 bp in length
1062	1161:	gap of 100 bp
1162	2202:	contig of 1041 bp in length
2203	2302:	gap of 100 bp
2303	3358:	contig of 1056 bp in length
3359	3458:	gap of 100 bp
3459	4520:	contig of 1062 bp in length
4521	4620:	gap of 100 bp
4621	5725:	contig of 1105 bp in length
5726	5825:	gap of 100 bp
5826	6895:	contig of 1070 bp in length
6896	6995:	gap of 100 bp
6996	8043:	contig of 1048 bp in length
8044	8143:	gap of 100 bp
8144	9212:	contig of 1069 bp in length
9213	9312:	gap of 100 bp
9313	10439:	contig of 1127 bp in length
10440	10539:	gap of 100 bp
10540	11612:	contig of 1073 bp in length
11613	11712:	gap of 100 bp
11713	12750:	contig of 1038 bp in length
12751	12850:	gap of 100 bp
12851	13865:	contig of 1036 bp in length
13867	13986:	gap of 100 bp
13987	15027:	contig of 1041 bp in length
15028	15127:	gap of 100 bp
15128	16197:	contig of 1070 bp in length
16198	16297:	gap of 100 bp
16298	17398:	contig of 1101 bp in length
17399	17498:	gap of 100 bp
17499	18610:	contig of 1112 bp in length
18611	18710:	gap of 100 bp
18711	19754:	contig of 1044 bp in length
19755	19854:	gap of 100 bp
19855	20944:	contig of 1090 bp in length
20945	21044:	gap of 100 bp
21045	22044:	contig of 1000 bp in length
22045	22144:	gap of 100 bp
22145	23221:	contig of 1077 bp in length
23222	23321:	gap of 100 bp
23322	24404:	contig of 1083 bp in length
24405	24504:	gap of 100 bp
24505	25643:	contig of 1139 bp in length
25644	25743:	gap of 100 bp
25744	26860:	contig of 1117 bp in length
26861	26960:	gap of 100 bp
26961	28028:	contig of 1068 bp in length
28029	28128:	gap of 100 bp
28129	29216:	contig of 1088 bp in length
29217	29316:	gap of 100 bp
29317	30393:	contig of 1077 bp in length


```

* 30394 30493: gap of 100 bp
* 30494 31584: contig of 1091 bp in length
* 31584 31684: gap of 100 bp
* 31684 32761: contig of 1077 bp in length
* 32761 32861: gap of 100 bp
* 32861 33951: contig of 1090 bp in length
* 33951 34051: gap of 100 bp
* 34051 35076: contig of 1024 bp in length
* 35076 35176: gap of 100 bp
* 35176 36272: contig of 1097 bp in length
* 36272 36372: gap of 100 bp
* 36372 37441: contig of 1069 bp in length
* 37441 37541: gap of 100 bp
* 37541 38600: contig of 1059 bp in length
* 38600 38700: gap of 100 bp
* 38700 39781: contig of 1081 bp in length
* 39781 39881: gap of 100 bp
* 39881 40979: contig of 1098 bp in length
* 40979 41079: gap of 100 bp
* 41079 42096: contig of 1017 bp in length
* 42096 42196: gap of 100 bp
* 42196 43295: contig of 1099 bp in length
* 43295 43395: gap of 100 bp
* 43395 44461: contig of 1066 bp in length
* 44461 44561: gap of 100 bp
* 44561 45625: contig of 1064 bp in length
* 45625 45725: gap of 100 bp
* 45725 46830: contig of 1105 bp in length
* 46830 46930: gap of 100 bp
* 46930 48033: contig of 1103 bp in length
* 48033 48133: gap of 100 bp
* 48133 49181: contig of 1048 bp in length
* 49181 49281: gap of 100 bp
* 49281 50381: contig of 1100 bp in length
* 50381 50481: gap of 100 bp
* 50481 51600: contig of 1119 bp in length
* 51600 51700: gap of 100 bp
* 51700 52817: contig of 1117 bp in length
* 52817 52917: gap of 100 bp
* 52917 53967: contig of 1050 bp in length
* 53967 54067: gap of 100 bp
* 54067 55103: contig of 1036 bp in length
* 55103 55203: gap of 100 bp
* 55203 56264: contig of 1061 bp in length
* 56264 56364: gap of 100 bp
* 56364 57469: contig of 1105 bp in length
* 57469 57569: gap of 100 bp
* 57569 58636: contig of 1067 bp in length
* 58636 58736: gap of 100 bp
* 58736 59797: contig of 1061 bp in length
* 59797 59897: gap of 100 bp
* 59897 60979: contig of 1082 bp in length
* 60979 61079: gap of 100 bp
* 61079 62035: contig of 956 bp in length
* 62035 62135: gap of 100 bp
* 62135 63164: contig of 1029 bp in length
* 63164 63264: gap of 100 bp
* 63264 64305: contig of 1041 bp in length
* 64305 64405: gap of 100 bp
* 64405 65427: contig of 1022 bp in length
* 65427 65527: gap of 100 bp
* 65527 66574: contig of 1047 bp in length
* 66574 66674: gap of 100 bp
* 66674 67683: contig of 1009 bp in length
* 67683 67783: gap of 100 bp
* 67783 68876: contig of 1093 bp in length
* 68876 68976: gap of 100 bp
* 68976 69998: contig of 1022 bp in length
* 69998 70098: gap of 100 bp
* 70098 71149: contig of 1051 bp in length
* 71149 71249: gap of 100 bp
* 71249 72339: contig of 1090 bp in length
* 72339 72439: gap of 100 bp

```

```

* 72440 73511: contig of 1072 bp in length
* 73511 73611: gap of 100 bp
* 73611 74726: contig of 1115 bp in length
* 74726 74826: gap of 100 bp
* 74826 75864: contig of 1038 bp in length
* 75864 75964: gap of 100 bp
* 75964 77015: contig of 1051 bp in length
* 77015 77115: gap of 100 bp
* 77115 78183: contig of 1068 bp in length
* 78183 78283: gap of 100 bp
* 78283 79298: contig of 1015 bp in length
* 79298 79398: gap of 100 bp
* 79398 80451: contig of 1053 bp in length
* 80451 80551: gap of 100 bp

Query Match      2.0% Score 55; DB 2; Length 81667;
Best Local Similarity 100.0%; Pred. No. 6.7e-19; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 0;

Qy      1561 TGTGGACCGCTTAAAGTACTTCAATTCCTAAGAGACTTGATGATTTA 1615
Db      4699 TGTGGACCGCTTAAAGTACTTCAATTCCTAAGAGACTTGATGATTTA 4753

RESULT 22
AC095790
LOCUS
DEFINITION
Rattus norvegicus clone CH230-9M12, *** SEQUENCING IN PROGRESS ***
AC095790
VERSION
AC095790.4 GI:22773054
KEYWORDS
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 179667)
Muzny D,Marie, Metzker M, Lee, Abramson S, Adams C, Alder J,
Allen C, Allen H, Alsbrooks S, Amin A, Angiano D,
Ayalebech V, Aoyagi A, Ayodeji M, Baca E, Baden H,
Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F,
Biswal K, Blair K, Blankenburg K, Blyth P, Brown M,
Bryant N, Buhay C, Burch P, Burrell K, Calderon E,
Cardenas V, Carter K, Cavazos I, Cesari H, Center A,
Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z, Chu J,
Cleveland C, Cockrell R, Cox C, Coyte M, Cree A, D'Souza L,
Davila M, Davis C, Davy-Carroll J, De Andrade C, Dederich D,
Delgado O, Denson S, Deramo C, Ding Y, Dinh H, Divya K,
Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K,
Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G,
Fernandez S, Finley M, Flagg N, Forbes L, Foster M, Foster P,
Frazer C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M,
Gebregorgis E, Geer K, Gill R, Grady M, Guerra W, Guevara M,
Gunnarsson P, Haaland W, Hamill C, Hamilton C, Hamilton K,
Harvey Y, Hawlak P, Hawes A, Henderson N, Hernandez J,
Hernandez R, Hines S, Hladun S, Hodge S, Idlebird D, Jackson A,
Hollins B, Howells S, Hulky S, Hume J, Idlebird D, Jackson A,
Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolivet A,
Karpathy S, Kelly S, Kelly S, Khan Z, King L, Kovar C,
Kowals C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J,
Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J,
Lorenshewa L, Louissege H, Losado R, Lu X, Ma J,
Maheshwari M, Mahindaratne M, Mahmoud M, Malloy K, Mangum A,
Mangum B, Mapua P, Martin K, Martin R, Martinez E,
Mawhinney S, McLeod M, McNeill T, Meenen E,
Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S,
Morgan M, Morris K, Morris S, Mundasa M, Murphy M, Nair L,
Nankervis C, Neal D, Newton N, Nguyen N, Norris S,
Nwaekelam O, Okwom G, Olarunsaogun A, Pal S, Parks K,
Paeternak S, Paul H, Perez A, Perez L, Pflanzsch C,
Plopper F, Poindester A, Popovic D, Primus B, Pu L,
Puzo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Rivels,B., Rellly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Sanders,M., Saverly,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sned,A., Sodergren,E., Song,X.-Z., Sorrell,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,D., Waldron,L., Walker,B., Wang,J.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 179667)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 179667)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (03-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 10, 2002 this sequence version replaced gi:21293770.
The sequence in this assembly is a combination of BAC-based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDJW
Center clone name: CH230-9M12
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 170329 bases at least Q40
Consensus quality: 171740 bases at least Q30
Consensus quality: 172810 bases at least Q20
Estimated insert size: 192053; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 89981: contig of 89981 bp in length
*
* 89982 90081: gap of unknown length
*
* 90082 179667: contig of 89586 bp in length.
Location/Qualifiers
1. 179667
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9M12"
90082. .91885
/note="wgs_contig"

```

Query Match	1.8%	Score 50	DB 2	Length 179667
Best Local Similarity	100.0%	Pred. No. 4,6e-16		
Matches	50	Conservative 0	Mismatches 0	Indels 0
Gap	2291	AGCCGAGAGGCTCTGTGTACAGACTTTCATGTGACGTCATCAACGACCCCA	2340	
Db	158228	AGCCGAGAGGCTCTGTGTACAGACTTTCATGTGACGTCATCAACGACCCCA	158277	
RESULT 23				
AC123888		273260 bp	DNA	linear
LOCUS				HTG 15-NOV-2002
DEFINITION				working draft sequence, 3
AC123888				
AC123888.4		GI:25007955		
KEYWORDS		HTG, HTGS, PHASE1, HTGS, DRAFT, HTGS, FULLTOP.		
SYNOPSIS		Rattus norvegicus (Norway rat)		
SOURCE		Rattus norvegicus		
ORGANISM		Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE		1 (bases 1 to 273260)		
AUTHORS		Murphy, D., Martie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anylebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, D., Blankensbury, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, J., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, V., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., De Souza, L., Davila, M. L., Davis, C., Day, Carroll, L., De Anda, C., Dedrich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georgegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guetara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howell, S., Hulky, S., Hume, J., Idlebird, D., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, K., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzen, L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martineau, B., Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Narkervic, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokwelen, O., Okwum, G., Olariu, S., Olson, A., Pal, S., Parks, K., Paternak, S., Paul, H., Perez, A., Perez, L., Plambeck, C., Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quirio, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Sytek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von		

Niederhuser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 273260)
 Worley, K.C.
 Direct Submission
 Submitted (03-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 273260)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 15, 2002 this sequence version replaced gi:23267611.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYD
 Center clone name: CH230-43J18
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 22476 bases at least Q40
 Consensus quality: 224733 bases at least Q30
 Consensus quality: 226257 bases at least Q20
 Estimated insert size: 227079; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/doc/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 268389: contig of 268389 bp in length
 * 268390 268489: gap of unknown length
 * 268490 269678: contig of 1189 bp in length
 * 269679 273260: gap of unknown length
 * 269779 273260: contig of 3482 bp in length.
 Location/Qualifiers
 1..273260
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-43J18"
 1..1061
 /note="wgs_end_extension
 clone_end:77"
 43606..44451
 /note="clone_boundary
 clone_end:17
 site:

misc_feature
 misc_feature
 misc_feature

misc_feature end_sequence: BH265538"
 58743..60404
 /note="wgs_contig"
 misc_feature 265291..266193
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence: BH265541"
 43343 C 44553 G 72796 C 45914 others

BASE COUNT 66654 A 43343 C 44553 G 72796 T 45914 others
 ORIGIN

Query Match 1.8%; Score 50; DB 2; Length 273260;
 Best Local Similarity 100.0%; Pred. No. 4.3e-16;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2291 AGCCAGAGGTGTGTACAGAGACTTCATGACGATCATCAACACCCCA 2340
 |||||||
 Db 44201 ACCGAGAGGCTGTGTACAGACTTCATGACGATCATCAACACCCCA 44250

RESULT 24
 MUSC1
 LOCUS
 DEFINITION Mus musculus sema3a mRNA for semaphorin 3A, complete cds.
 ACCESSION D85028
 VERSION D85028.2 GI:18389545
 KEYWORDS collapse-1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2356)
 Taniguchi, M., Yuasa, S., Fujisawa, H., Naruse, I., Saga, S., Mishina, M.
 and Yagi, T.
 Disruption of semaphorin III/D gene causes severe abnormality in
 peripheral nerve projection
 Neuron 19 (3), 519-530 (1997)
 MEDLINE 97470885
 PUBMED 9331345
 REFERENCE 2 (bases 1 to 5952)
 Taniguchi, M.
 Direct Submission
 Submitted (02-MAY-1996) Masahiko Taniguchi, Department of
 Biochemistry and Molecular Biology, Graduate School of Medicine,
 The University of Tokyo; Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033,
 Japan (E-mail: taniguchi@m.u-tokyo.ac.jp, Tel: 81-3-5802-2925,
 Fax: 81-3-3813-8732)
 On Jan 28, 2002 this sequence version replaced gi:1313903.

COMMENT
 FEATURES
 source
 1..5952
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..5952
 /gene="sema3A"
 /gene="sema3A"
 652..2970
 /codon_start=1
 /product="semaphorin 3A"
 /protein_id="BA19773.1"
 /db_xref="GI:2055316"
 /translation="MGWFTGACLFMGVLLTARANYANKNNVRLKLSYKEMLESNN
 VITNGLANSSSYHTFLIDERSRLVYGAOKHIFPNLVNIDPOKIVPVSYPYRDE
 CKMAGKOLKEKCANFIVLEAYNOTHLYACGTGAFFICTYIEVGHHEVDI FKLQDS
 HENGGRGKSPYEDPLLTASLLIDELVSGTADMGDPALFRLGHHHPYRTOHDS
 RMNDPRTSAHLIPESDNPDDKYTFPFRFNALDGRSGKATYARIGQIKNDPGGH
 RSLVKNKWTFLKARLCSVPGNGIDHFDLQVFLNSKDPKNPVLYGVFTTSSNI
 FSGSAVCMYSMSDVRVFLGYPARHDPNQWYQGVPRPGTCSKTFGGFDS
 KDLPDVITFARSHAPAMNPFVPIINBPIMIKTVNYQFTQIVDRAVDAEDGQYVW
 IGTGVTVLKVSVKPEKTHDLEVLLEBMVQVLEPPTISAMELSTKQOOLYIGSDG
 VAQLPIRCDIYKACACCAARDPYCAMOSSCSRYPRPKRTRRBDIRNGPLTH
 CSDIQHDNHGSPLEERITLYGVNSSTFLFECSSKSGSRALVYVWQPRNBRKEIRK
 GDHITRTQGLLRSLQKDSGNYLCAVHEGFMOTLLKTVLEVIDTEHLEELHKOD

DDGSKIKEMSSMTPTSOQKWRVDFMOLINHNNTMDFECCQWKRKORORPGH
 SGGSSNKKKHQOESKGRNRRTHFERAPRSV"

BASE COUNT 1829 a 1203 c 1214 g 1706 t

ORIGIN

Query Match 1.7%; Score 47; DB 10; Length 5952;
 Best Local Similarity 100.0%; Pred. No. 2.9e-14;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 164
 |||||
 570 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 616

RESULT 25
 AC068065/c
 LOCUS
 DEFINITION Mus musculus clone RP23-30905 strain C57BL6/J, WORKING DRAFT
 AC068065
 AC068065.1 GI:7658325
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 195335)
 Huang, M.C., Idol, J., Lee-Lin, S.-Q., Maduro, Q.L., Maduro, V.B.,
 Maestrian, S.D., McCloskey, J.C., Morse, J.W., Ojodu, M.A., Pearson, R.,
 Stantrop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
 Tiongsen, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
 Wetherby, K.D. and Green, E.D.
 NISC Mouse Sequencing Initiative
 Unpublished
 2 (bases 1 to 195335)
 Green, E.D.
 TITLE Direct Submission
 AUTHORS Submitted (28-Apr-2000) NIH Intramural Sequencing Center, 8717
 JOURNAL Groveomont Circle, Gaithersburg, MD 20877, USA
 COMMENT

 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.mouse@nih.gov

 Project Information
 Center project name: YV
 Center clone name: 309J05

 Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 187134 bases at least Q40
 Consensus quality: 188869 bases at least Q30
 Consensus quality: 190074 bases at least Q20
 Insert size: 19000; agarose-fp
 Insert size: 19400; pulse-field-gel
 Insert size: 195335; sum-of-contigs
 Quality coverage: 8.15x in Q20 bases; agarose-fp
 Quality coverage: 7.98x in Q20 bases; pulse-field-gel
 Quality coverage: 7.92x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 2293: contig of 2293 bp in length
 gap of unknown length

2294 5070: contig of 2777 bp in length
 * gap of unknown length
 * 5071 8818: contig of 3748 bp in length
 * gap of unknown length
 * 8819 16171: contig of 7353 bp in length
 * gap of unknown length
 * 16172 30056: contig of 13885 bp in length
 * gap of unknown length
 * 30057 47096: contig of 17040 bp in length
 * gap of unknown length
 * 47097 68618: contig of 21522 bp in length
 * gap of unknown length
 * 68619 95138: contig of 26520 bp in length
 * gap of unknown length
 * 95139 126583: contig of 31445 bp in length
 * gap of unknown length
 * 126584 156528: contig of 22945 bp in length
 * gap of unknown length
 * 156529 195335: contig of 38807 bp in length.
 Location/Qualifiers
 1..195335
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL6/J"
 /db_xref="taxon:10090"
 /clone="RP23-30905"
 /clone_lib="RP23-30905" mouse BAC library 23"
 BASE COUNT 61274 a 35333 c 35501 g 62857 t 370 others

ORIGIN

Query Match 1.7%; Score 47; DB 2; Length 195335;
 Best Local Similarity 100.0%; Pred. No. 2.4e-14;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 164
 |||||
 27069 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 27023

RESULT 26
 AC121125/c
 LOCUS
 DEFINITION Mus musculus clone RP23-10F17, *** SEQUENCING IN PROGRESS ***
 AC121125
 AC121125.4 GI:31581674
 VERSION HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 265982)
 Birren, B., Nussbaum, C. and Lander, E.
 Mus musculus, Clone RP23-10F17
 Unpublished
 2 (bases 1 to 265982)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Banna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Riley, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Labrecque, K., Lamazares, R.,
 Landers, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,

VERSION L41541.1 GI:18654369
 KEYWORDS collapsin/semaphorinIII.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2319)
 Kimura, T. and Fishman, M.C.
 cDNA sequence of mouse Collapsin/SemaphorinIII
 JOURNAL Unpublished
 FEATURES
 location/Qualifiers
 1..2319
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /dev_stage="adult"
 1..2319
 /note="putative"
 /codon_start=1
 /product="collapsin/semaphorin III"
 /protein_id="AA177611.1"
 /db_xref="GI:18654370"
 /translation="MGWFTGIACLFMGVLLTARANYANGKNVPRILKSYKMEESNN
 VITFNGLANSSYHTFLDEERSRLVYGAQKHIFSFLVNIKDFOKLVMPVSYTRDS
 CKWAGKDLKECANFIKYLEAVNQTHTYAGCTGAHPICVYIEVGHHPDENITFLQDS
 HFENGKRSYDPRFLTASLLIDGELYSCTAADPWGRDPAIFRTLDGHHPIRTQDS
 RMLNDPRFISAHLLPESDNPEDDKVYFFEFENAIIGESHGKATHARIGQICXNDPGGH
 RSLVNMKTTFLKARLICVSPGNGIDTHPELDVFLMNSKDPKNPIYGVFTSSNI
 PKSAVCMYSMDVRVPLGYPYARHDGPNYQVYQGVPPYPRGTCSKTFGGFDS
 KDLDDVITFAKSHPMNVPFPIINRPIMIKTDVNYQFOIIVDRAEDGQYDVMF
 IGTGVNVLKVVSVPKETWDLLEVLLEMTVFEPTTISAMELSTKQOQYIGSTAG
 VAOLPLHRCDIYGAACCECLARDPYCAMDSSCSRYFPFAKRTKODIRNGDLTH
 CSDLQHDNHHGSPLEERIIYGVENSSTFLSCPSQALVYQFORNRRRSKEIEM
 GDHIIIRTEQGLIRSLQKDSGNVLCVAHGHGEMOTLLKTLVETDIEHLEELHKD
 DGGSKIKEMSSSTPSOKVYWRDPMOLINPNLNTMDEFCQYWKDRKORRPPGH
 SQSSNMKMHQSKGKRNRTHEPERAPRSV"
 BASE COUNT 703 a 513 c 549 g 554 t
 ORIGIN
 Query Match 1.5%; Score 41; DB 10; Length 2319;
 Best Local Similarity 100.0%; Pred. No. 8.1e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2171 CAAACTCTTTTAAAGTAACCTGGAAGTCATTGACACAGA 2211
 |||||
 Db 1975 CAAACTCTTTTAAAGTAACCTGGAAGTCATTGACACAGA 2015
 |||||
 RESULT 29
 MMRNASEMD 2913 bp mRNA linear ROD 08-JUL-1996
 LOCUS Mus musculus mRNA for semaphorin D.
 DEFINITION X85993
 ACCESSION X85993.1 GI:854329
 VERSION X85993.1
 KEYWORDS semaphorin; semd gene.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Puschel, A.W., Adams, R.H. and Betz, H.
 Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension
 JOURNAL Neuron 14 (5), 941-948 (1995)
 MEDLINE 95267431
 PUBMED 7748561
 2 (bases 1 to 2913)
 Adams, R.H.
 Direct Submision
 Submitted (30-MAR-1995) R.H. Adams, Max-Planck-Institute fuer
 Hirnforschung, Deutscherordenstr. 46, D- 60528 Frankfurt, FRG
 Sequence overlapping with those under the acc#U02528, L26080,
 L26081 & L26082.

FEATURES
 source
 Location/Qualifiers
 1..2913
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone_1db="Stratagene lambda ZAP"
 /dev_stage="12 days embryo"
 1..2913
 /gene="semd"
 111..2429
 /gene="semd"
 /codon_start=1
 /product="semaphorin D"
 /protein_id="CAA59985.1"
 /db_xref="GI:854330"
 /db_xref="MG1:107558"
 /translation="MGWFTGIACLFMGVLLTARANYANGKNVPRILKSYKMEESNN
 VITFNGLANSSYHTFLDEERSRLVYGAQKHIFSFLVNIKDFOKLVMPVSYTRDS
 CKWAGKDLKECANFIKYLEAVNQTHTYAGCTGAHPICVYIEVGHHPDENITFLQDS
 HFENGKRSYDPRFLTASLLIDGELYSCTAADPWGRDPAIFRTLDGHHPIRTQDS
 RMLNDPRFISAHLLPESDNPEDDKVYFFEFENAIIGESHGKATHARIGQICXNDPGGH
 RSLVNMKTTFLKARLICVSPGNGIDTHPELDVFLMNSKDPKNPIYGVFTSSNI
 PKSAVCMYSMDVRVPLGYPYARHDGPNYQVYQGVPPYPRGTCSKTFGGFDS
 KDLDDVITFAKSHPMNVPFPIINRPIMIKTDVNYQFOIIVDRAEDGQYDVMF
 IGTGVNVLKVVSVPKETWDLLEVLLEMTVFEPTTISAMELSTKQOQYIGSTAG
 VAOLPLHRCDIYGAACCECLARDPYCAMDSSCSRYFPFAKRTKODIRNGDLTH
 CSDLQHDNHHGSPLEERIIYGVENSSTFLSCPSQALVYQFORNRRRSKEIEM
 GDHIIIRTEQGLIRSLQKDSGNVLCVAHGHGEMOTLLKTLVETDIEHLEELHKD
 DGGSKIKEMSSSTPSOKVYWRDPMOLINPNLNTMDEFCQYWKDRKORRPPGH
 SQSSNMKMHQSKGKRNRTHEPERAPRSV"
 BASE COUNT 888 a 632 c 679 g 714 t
 ORIGIN
 Query Match 1.5%; Score 41; DB 10; Length 2913;
 Best Local Similarity 100.0%; Pred. No. 8e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2171 CAAACTCTTTTAAAGTAACCTGGAAGTCATTGACACAGA 2211
 |||||
 Db 2085 CAAACTCTTTTAAAGTAACCTGGAAGTCATTGACACAGA 2125
 |||||
 RESULT 30
 AC109862_1/c
 WPCOMMENT
 Sequence split into 4 fragments LOCUS AC109862 Accession AC109862
 Fragment Name Begin End
 AC109862_1 1 110000
 AC109862_2 100001 210000
 AC109862_3 200001 310000
 AC109862_4 300001 381240
 Continuation (2 of 4) of AC109862 from base 100001 (AC109862 Rattus norvegicus clone CH2).
 Query Match 1.5%; Score 41; DB 2; Length 110000;
 Best Local Similarity 100.0%; Pred. No. 6.5e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 569 TATATCAGACTCACTGTGACGCTGTGAACGGGGCTTT 609
 |||||
 Db 99404 TATATCAGACTCACTGTGACGCTGTGAACGGGGCTTT 99364
 |||||
 RESULT 31
 AC102452 184006 bp DNA linear HTG 09-JUN-2003
 LOCUS Mus musculus clone RP24-272120, *** SEQUENCING IN PROGRESS ***
 DEFINITION ordered pieces.
 ACCESSION AC102452
 VERSION AC102452.5 GI:31544113
 KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 184006)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP24-272L20
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 184006)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B.,
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
 Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hages, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K.,
 Lamezates, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Margolis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
 Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trisillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 184006)
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choquel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearlano, K.,
 Diaz, J., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hages, B., Hagopian, D., Hages, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C.,
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (09-JUN-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jun 9, 2003 this sequence version replaced gi:11245823.
 All repeats were identified using RepeatMasker:
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: 118837
 Center clone name: 272_L_20

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 66973: contig of 66973 bp in length
 * 66974 67073: gap of 100 bp
 * 67074 159784: contig of 92711 bp in length
 * 159785 159884: gap of 100 bp
 * 159885 184006: contig of 24122 bp in length.
 Location/Qualifiers
 1..184006
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-272L20"
 /clone_11b="RP24-272L20"
 /clone_11b="RP24-272L20" Male Mouse BAC"
 57869 a 34540 c 34334 g 57056 t 207 others

BASE COUNT 57869 a 34540 c 34334 g 57056 t 207 others
 ORIGIN

Query Match 1.5% Score 41; DB 2; Length 184006;
 Best Local Similarity 100.0%; Pred. No. 6.3e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2171 CAACTCTTCTTAAGTAACCTGGAAGTCAATGACACAGA 2211
 Db 21139 CAACTCTTCTTAAGTAACCTGGAAGTCAATGACACAGA 21179

RESULT 32
 AC022368/c 207757 bp DNA linear ROD 30-JAN-2001
 LOCUS AC022368/c
 DEFINITION Mus musculus clone RP23-192D21 strain C57BL/6/J, complete sequence.
 ACCESSION AC022368
 VERSION AC022368.2 GI:12597739
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 207757)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE Mus musculus, clone RP23-192D21

TITLE Direct Submission
 JOURNAL Submitted (02-FEB-2000) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 REFERENCE 3 (bases 1 to 207757)
 AUTHORS Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (30-JAN-2001) NIH Intramural Sequencing Center, 8717
 Groveont Circle, Gaithersburg, MD 20877, USA
 COMMENT On Jan 30, 2001 this sequence version replaced gi:16855198.

Center: NIH Intramural Sequencing Center
 Web site: http://www.nisc.nih.gov
 Contact: nisc.mouse@nih.gov
 Project Information
 Center project name: 192D21
 Center clone name: 192D21

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the features section.

FEATURES

SOURCE

1..207757
location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-192D21"
/clone_lib="RPCI mouse BAC library 23"
32117..32123
/note="low quality single stranded/single chemistry
region"
misc_feature
32123..32166
/note="single clone coverage"
misc_feature
101083..102298
/note="single clone coverage"
misc_feature
102471..102703
/note="single clone coverage"
misc_feature
102486..102493
/note="low quality single stranded/single chemistry
region"
misc_feature
151895..152199
/note="single clone coverage"
misc_feature
162474..162566
/note="single clone coverage"
misc_feature
190297..190474
/note="single clone coverage"
BASE COUNT 65923 a 38017 c 37853 g 65964 t
ORIGIN

Query Match 1.5%; Score 41; DB 10; Length 207757;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OR 2171 CAAACTCTTCTTAAGTAACTGGAAGTCATTGACACAGA 2211
Db 77684 CAACTCTTCTTAAGTAACTGGAAGTCATTGACACAGA 77644

RESULT 33
AC068065 195335 bp DNA linear HTG 28-APR-2000
LOCUS Mus musculus clone RP23-309J5 strain C57BL6/J, WORKING DRAFT
DEFINITION
SEQUENCE 11 unordered pieces.
AC068065
AC068065.1 GI:7658325
VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195335)

Beckstrom-Stenberg, S.M., Benjamin, B., Blakeley, R.W.,
Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,
Huang, M.C., Idol, J., Lee-Ihn, S.-Q., Maduro, Q.L., Maduro, V.B.,
Mastrian, S.D., McCloskey, J.C., Morse, E., Ojodu, M.A., Pearson, R.,
Stancirpop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
Wetherby, K.D. and Green, E.D.
TITLE NISC Mouse Sequencing Initiative
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 195335)
AUTHORS Green, E.D.
TITLE Direct SubMISSION
JOURNAL Submitted (28-APR-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nhgri.nih.gov
Project Information
Center project name: YV
Center clone name: 309J05

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187134 bases at least Q40
Consensus quality: 18869 bases at least Q30
Consensus quality: 190074 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 194000; pulse-field-gel
Insert size: 195335; sum-of-contigs
Quality coverage: 8.15x in Q20 bases; agarose-fp
Quality coverage: 7.98x in Q20 bases; pulse-field-gel
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2293: contig of 2293 bp in length
* 2294 5070: contig of 2777 bp in length
* 5071 8818: contig of 3748 bp in length
* 8819 16171: contig of 7353 bp in length
* 16172 30056: contig of 13885 bp in length
* 30057 47096: contig of 17040 bp in length
* 47097 68618: contig of 21522 bp in length
* 68619 95138: contig of 26520 bp in length
* 95139 126583: contig of 31445 bp in length
* 126584 156528: contig of 29945 bp in length
* 156529 195335: contig of 38807 bp in length.

FEATURES

SOURCE

1..195335
location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-309J5"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 61274 a 35333 c 35501 g 62857 t 370 others
ORIGIN

Query Match 1.2%; Score 33; DB 2; Length 195335;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OR 310 AGAATGTTGATCAACATGATGATCACTT 342
|||||

DB 176419 AGAATGTTGATCCACATGATGATCTTT 176451

RESULT 34

AC121585 195311 bp DNA linear ROD 14-DEC-2002

LOCUS Mus musculus chromosome 5 clone RP23-265M4, complete sequence.

DEFINITION AC121585

AC121585.3 GI:26801335

HTG.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 195311)

AUTHORS McPherson,J.D. and Waterston,R.H.

TITLE The sequence of Mus musculus clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 195311)

AUTHORS McPherson,J.D. and Waterston,R.H.

TITLE Direct Submision

JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 195311)

AUTHORS McPherson,J.D. and Waterston,R.H.

TITLE Direct Submision

JOURNAL Submitted (04-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 195311)

AUTHORS McPherson,J.D. and Waterston,R.H.

TITLE Direct Submision

JOURNAL Submitted (14-DEC-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT On Dec 14, 2002 this sequence version replaced gi:22711832.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

Contact: submissions@wustl.wustl.edu

----- Project Information -----

Center project name: M_BA0265M04

----- Location/Qualifiers -----

1.195311

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="5"

/clone="RP23-265M4"

BASE COUNT 63171 a 38828 c 37957 g 55355 t

ORIGIN

Query Match 1.2%; Score 32; DB 10; Length 195311;

Best Local Similarity 100.0%; Pred. No. 8.6e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 GAAATGTTGATCCACATGATGATCTTT 342

DB 85566 GAAATGTTGATCCACATGATGATCTTT 85587

RESULT 35

AX359936/c 4661 bp DNA linear PAT 13-FEB-2002

LOCUS Sequence 10 from Patent WO0202783.

DEFINITION AX359936

AX359936.1 GI:18675575

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 4661)

AUTHORS Miller,A.D.

TITLE Improved retroviral vectors for gene transfer and expression

JOURNAL Biotechniques 7 (9), 980-982 (1989)

AUTHORS Bleck,G.T.

TITLE Expression vectors

JOURNAL Patent: WO 0202783-A 10 10-JAN-2002;

Galat Design, Inc. (US)

FEATURES

source

1.4661

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

/note="synthetic"

BASE COUNT 1257 a 1096 c 1161 g 1147 t

ORIGIN

Query Match 1.1%; Score 31; DB 6; Length 4661;

Best Local Similarity 100.0%; Pred. No. 4e-05;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGTTAACACACTT 31

DB 4041 AATCTTTATTTATGATGTTAACACACTT 4011

RESULT 36

AX382150/c 4661 bp DNA linear PAT 18-MAR-2002

LOCUS Sequence 10 from Patent WO0202738.

DEFINITION AX382150

AX382150.1 GI:19576959

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 4661)

AUTHORS Brenzel,R.D., Miller,L.U., Bleck,G.T. and York,D.

TITLE Host cells containing multiple integrating vectors

JOURNAL Patent: WO 0202738-A 10 10-JAN-2002;

Galat Design, Inc. (US)

FEATURES

source

1.4661

/organism="synthetic construct"

/mol_type="genomic DNA"

/db_xref="taxon:32630"

BASE COUNT 1257 a 1096 c 1161 g 1147 t

ORIGIN

Query Match 1.1%; Score 31; DB 6; Length 4661;

Best Local Similarity 100.0%; Pred. No. 4e-05;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGTTAACACACTT 31

DB 4041 AATCTTTATTTATGATGTTAACACACTT 4011

RESULT 37

SYMMPLP1N3 6620 bp DNA linear SYN 05-JAN-2001

LOCUS Retroviral vector pLNCX, complete sequence.

DEFINITION M28247

M28247.1 GI:208846

KEYWORDS genome; neomycin phosphotransferase; retroviral vector.

SOURCE Retroviral vector pLNCX

ORGANISM

REFERENCE 1 (bases 1 to 6620)

AUTHORS Miller,A.D.

TITLE Direct Submision

JOURNAL Submitted (22-SEP-1989) A.D. Miller, Program in Molecular Medicine, Fred Hutchinson Cancer Research Center, Seattle, WA 98104, USA

REFERENCE 2 (bases 1 to 6620)

AUTHORS Miller,A.D. and Rosman,G.J.

TITLE Improved retroviral vectors for gene transfer and expression

JOURNAL Biotechniques 7 (9), 980-982 (1989)

FEATURES
MEDLINE 90211945
PUBMED 2631796
source
location/Qualifiers
1. .6620
/organism="Retroviral vector pLNCX"
/mol_type="genomic DNA"
/db_xref="taxon:147133"
/note="derived from Moloney murine leukemia virus"
1447145
/note="mouse DNA end-Mo-MuSV DNA start"
145. .733
/note="Mo-MuSV 5' long terminal repeat"
803. .1612
/note="extended packaging signal"
113371134
/note="Mo-MuSV DNA end-Mo-MuLV DNA start"
1193. .1195
/note="gag ATG start codon to TAG stop codon"
161671617
/note="Mo-MuLV DNA end-Tn5 DNA start"
1656. .2450
/note="neomycin phosphotransferase"
/codon_start=1
/transl_table=1
/protein_id="AA72064.1"
/db_xref="GI:208847"
/translation="MIEQDGLHAGSPAAWVERLFGYDMAOQTICSDAAVPRLSAQR
PVEFKTDISALNELDEARLSMTATGPGCAVLDVTEARDMILLEGVGGDL
LSHLAPAEKVSIMADAKRLHTLDPCPDHQAKHIERAKTRMEGLVDQDLDB
EHQGLAPAELEFARLKARMPDEDLVTHGDCLEINWENGRFSGFTIDCGRLGVADRY
QDIALATRDIAEELEGWADRFVLVYGIAPDSORIAFYRLDEFF"
279972800
misc_recomb
/note="Tn5 DNA end-human cytomegalovirus DNA start"
2800. .3617
/note="human cytomegalovirus immediate early promoter"
361773618
/note="human cytomegalovirus DNA end-Mo-MuLV DNA start"
3666. .4259
/note="Mo-MuLV 3' long terminal repeat"
432874329
/note="Mo-MuLV DNA end-plasmid pBR322 DNA start"
BASE COUNT 1528 a 1807 c 1692 g 1593 t
ORIGIN
Query Match 1.1%; Score 31; DB 12; Length 6620;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 AATCTTTATTTATCGATGTTAACAAGCTT 31
|||||
Db 3639 AATCTTTATTTATCGATGTTAACAAGCTT 3609
RESULT 38
SYNMOV2/6
LOCUS 6837 bp DNA linear SYN 27-APR-1993
DEFINITION Moloney murine leukemia virus retroviral vector pLNCX, complete
sequence.
ACCESSION M64754
VERSION M64754.1 GI:208857
KEYWORDS cloning vector; histidinol dehydrogenase; retroviral vector.
SOURCE unidentified cloning vector
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6837)
AUTHORS Stockschlaeder, M.A., Storb, R., Osborne, W.R. and Miller, A.D.
TITLE L-histidinol provides effective selection of
retrovirus-vector-transduced keratinocytes without impairing their
proliferative potential
Hum. Gene Ther. 2 (1), 33-39 (1991)
JOURNAL Hum. Gene Ther. 2 (1), 33-39 (1991)
MEDLINE 91322161
PUBMED 1650586
COMMENT Original source, text: Synthetic, DNA.

FEATURES
source
location/Qualifiers
1. .6837
/organism="unidentified cloning vector"
/mol_type="genomic DNA"
/db_xref="taxon:45196"
175. .763
/note="Mo-MuSV 5' long terminal repeat"
833. .1642
/note="extended packaging signal"
1223. .1225
/note="gag ATG start codon to TAG stop codon"
1672. .2976
/EC_number="1.1.1.23"
/codon_start=1
/transl_table=1
/product="histidinol dehydrogenase"
/protein_id="AA73024.1"
/db_xref="GI:208858"
/translation="MSFNTLIDMNSCSPEOQALITRPATISASDSITRTVSDILDNK
TRGDDALREYSAREDKTEVTALRTPEEIAAGARLSDELKOMTAIVAKNIEFHSAO
TLAPVDVETQPGVRCQOVTVPVSVGLITPGSAPLPESTVLMATPARIGCCOVVLC
SPPIADBITIYAAQLCGVGEIENVGGAQATIALAFSGESVPKDKIFPGNATFTEK
RQVSQRUDGAIDMPAGSEVLVYADSGATPDVVASDLSQAHEGPOSQVLTLPDAD
IARKVAEAVEROLEAPRADTARQALSASRLIVTKDIAQCVAISNOYGPBRLIQTEN
ARDLVDAITSAGSVFLGDMSPESAGDVASGTHVLPYGYATCSISGLADFOKRMVY
OELSKGFSALASTIETLAAERLTHAKNAVTLRVNALKEQA"
3017. .3834
promoter
/note="human cytomegalovirus early promoter"
3883. .4476
/note="Mo-MuLV 3' long terminal repeat"
BASE COUNT 1610 a 1877 c 1722 g 1628 t
ORIGIN
Query Match 1.1%; Score 31; DB 12; Length 6837;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 AATCTTTATTTATCGATGTTAACAAGCTT 31
|||||
Db 3856 AATCTTTATTTATCGATGTTAACAAGCTT 3826
RESULT 39
AR302096/c
LOCUS 7160 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 5 from patent US 6541197.
ACCESSION AR302096
VERSION AR302096.1 GI:31690130
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.
REFERENCE 1 (bases 1 to 7160)
AUTHORS Link, C.J., Jr., Levy, J.P., Wang, S. and Seregina, T.
TITLE Vehicles for stable transfer of green fluorescent protein gene and
methods of use for same
Patent: US 6541197-A 5 01-APR-2003;
FEATURES Location/Qualifiers
1. .7160
/organism="unknown"
BASE COUNT 1647 a 2009 c 1872 g 1632 t
ORIGIN
Query Match 1.1%; Score 31; DB 6; Length 7160;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 AATCTTTATTTATCGATGTTAACAAGCTT 31
|||||
Db 4179 AATCTTTATTTATCGATGTTAACAAGCTT 4149
RESULT 40

```

AR302097/c  AR302097  7235 bp  DNA  linear  PAT 12-JUN-2003
LOCUS       AR302097
DEFINITION  Sequence 6 from patent US 6541197.
ACCESSION   AR302097
VERSION     AR302097.1  GI:31690131
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 7235)
AUTHORS    Link,C.J., Jr., Levy,J.P., Wang,S. and Seregina,T.
TITLE      Vehicles for stable transfer of green fluorescent protein gene and
METHODS    method of use for same
JOURNAL     Patent: US 6541197-A 6 01-APR-2003;
FEATURES    Location/Qualifiers
source      1..7235
            /organism="Unknown"
BASE COUNT  1662 a 2088 c 1880 g 1605 t
ORIGIN
Query Match 1.1%; Score 31; DB 6; Length 7235;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATTCAGTGTAAACAAGCTT 31
Db 4254 AATCTTTATTTATTCAGTGTAAACAAGCTT 4224

RESULT 41
AX591134 29 bp DNA linear PAT 27-JAN-2003
LOCUS       AX591134
DEFINITION  Sequence 7 from Patent WO02086157.
ACCESSION   AX591134
VERSION     AX591134.1  GI:27949649
KEYWORDS    .
SOURCE      unidentified
ORGANISM    unidentified
REFERENCE   1
AUTHORS    Latil,A., Cusancenot,O. and Algarde-Genin,M.
TITLE      Use of semaphorin 3a for diagnosing and treating cancer, especially
JOURNAL     prostate cancer
          Patent: WO 02086157-A 7 31-OCT-2002;
          Urogene Societe anonyme (FR)
FEATURES    Location/Qualifiers
source      1..29
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
            /note="Amorce utilisee pour supprimer le codon stop de la
            traduction de l'ADNc de la semaphorine-3A (amorce 5')"
BASE COUNT  9 a 6 c 5 g 9 t
ORIGIN
Query Match 1.1%; Score 29; DB 6; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00073;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2155 GGAACATGGGTCATACAACTCTTTTA 2183
Db 1 GGAACATGGGTCATACAACTCTTTTA 29

RESULT 42
AF086761 3148 bp mRNA linear VRT 16-JUL-1999
LOCUS       AF086761
DEFINITION  Danio rerio semaphorin zla mRNA, complete cds.
ACCESSION   AF086761
VERSION     AF086761.1  GI:5499716
KEYWORDS    .
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 3148)
AUTHORS    Yee,C.S., Chandrasekhar,A., Halloran,M.C., Shoji,W., Warren,J.T.
            and Kuwada,J.Y.
TITLE      Molecular cloning, expression, and activity of zebrafish semaphorin
            zla
JOURNAL     Brain Res. Bull. 48 (6), 581-593 (1999)
MEDLINE     99313409
PUBMED      10386838
REFERENCE   2 (bases 1 to 3148)
AUTHORS    Yee,C.S., Chandrasekhar,A., Halloran,M.C., Shoji,W., Warren,J.T.
            and Kuwada,J.Y.
TITLE      Direct Submision
JOURNAL     Submitted (24-AUG-1998) Biology, University of Michigan, 830 N.
            University, Ann Arbor, MI 48109-1048, USA
FEATURES    Location/Qualifiers
source      1..3148
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /codon_start=1
            /product="semaphorin zla"
            /protein_id="AAD43964.1"
            /db_xref="GI:5499717"
            /translation="MDYLVGIFLLCGVALPGVAPQHTKENPRLKSYNEMLESSN
            LVTPLGLANSSGYDTFLMDGERGLLVGAEDHVSFDPD.VINRDKOIAWATPSKRD
            ECKAAGRLRDCNFRVAVLOSNTQTHYICGTAFPHICGSELEKRAENRIPIDA
            NYFENGKGSFLYDPMQSSLLDGEELSGTSADPFMGKDFALFRLGSHHRTIQHD
            SRWLNPERFLGIHLIPESDNPEDDKIFLFFENAMDEGHTKATISRGQCKDMG
            HRSLVNKKWTFELKAKLTCSPVGLANGIDIPBELDVFMSAKDPNPIYAVFTTSSN
            IFRSGALCMYMAIDIRVFLGPGVAPARDGPNQWVPFGORVYPPPGTCSPTGGFDS
            TKDLPDVITFRAPAMVNPVPMGKGIIVRTNVEYQFQVLVDRVREARDGGVDV
            FIGTDLGTVLKVITIPRESMIDLEVLIEKTVREPRITAMSELSTQOQLYIGSDL
            GISQMPHRCVYGAACCCCLARDPYCAMDTECSRYPFAKRTKRDITRNDPPLS
            QCSDLHNHNDLLEGYSVERSVYEVNSMFLSCSPKQALITWQLQKPNDEKHEI
            VIDRLSLTGGGLIRSLTQADSGVFLCHAVEGFIQPLRIINLQVIPSQVGEHLIR
            AGTNDKDPAPKRLMYRDFMSLLHPDLNVSDEFCERIMKREKKKGAQKRVNGT
            VSTNKKETPOTTAOSLONPTORANAPVPNPSPVQIFPKSGTGORSGSPGTVSTES
            OSTNPDYTKAESQRAQPNQAKKGPQTPORPPTAKKQIQENKRGNRRTHEQRP
            FRSV"
BASE COUNT  864 a 804 c 800 g 680 t
ORIGIN
Query Match 1.0%; Score 27; DB 5; Length 3148;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 500 GATGAATGCAAGTGGCTGGAAGAAGAC 526
Db 661 GATGAATGCAAGTGGCTGGAAGAAGAC 687

RESULT 43
AX107928 9729 bp DNA linear PAT 30-APR-2001
LOCUS       AX107928
DEFINITION  Sequence 56 from Patent WO0125466.
ACCESSION   AX107928
VERSION     AX107928.1  GI:13923311
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS    Slingsby,J., Kingsman,S.M., Rohll,J.O. and Slade,A.O.
TITLE      Producer cell for the production of retroviral vectors
JOURNAL     Patent: WO 0125466-A 56 12-APR-2001;
            Oxford Biomedica (UK) Limited (GB)
FEATURES    Location/Qualifiers
source      1..9729
            /organism="synthetic construct"

```

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic sequence"
BASE COUNT 2270 a 2657 c 2540 g 2262 t
ORIGIN

Query Match 1.0%; Score 27; DB 6; Length 9729;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTAACAA 27
|||||
Db 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 44
LOCUS AX107927 12473 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 55 from Patent WO0125466.
ACCESSION AX107927
VERSION AX107927.1 GI:13923310
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Slingsby, J., Kingsman, S.M., Rohlf, J.O. and Slade, A.O.
TITLE Producer cell for the production of retroviral vectors
JOURNAL Patent: WO 0125466-A 55 12-APR-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
SOURCE 1. 12473
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic sequence"

BASE COUNT 2841 a 3434 c 3357 g 2841 t
ORIGIN

Query Match 1.0%; Score 27; DB 6; Length 12473;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTAACAA 27
|||||
Db 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 45
LOCUS AX591129 25 bp DNA linear PAT 27-JAN-2003
DEFINITION Sequence 2 from Patent WO02086157.
ACCESSION AX591129
VERSION AX591129.1 GI:27949644
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Latil, A., Cussenot, O. and Algate-Genin, M.
TITLE Use of semaphorin 3a for diagnosing and treating cancer, especially prostate cancer
JOURNAL Patent: WO 02086157-A 2 31-OCT-2002;
Urogene Societe anonyme (FR)
FEATURES
SOURCE 1. 25
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="Amorce d'amplification du gene de la semaphorine-3a(lower primer) choisie avec l'assistance des programmes informatiques Oligo4 (National Biosciences, Plymouth, MN) et Primer express (Perkin-Elmer Applied

Bioystems, Foster city, CA)"
BASE COUNT 7 a 7 c 4 g 7 t
ORIGIN

Query Match 0.9%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1556 ACAGTGTGGACCGTTCTTAAG 1580
|||||
Db 25 ACAGTGTGGACCGTTCTTAAG 1

RESULT 46
LOCUS AX207158 1400 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 5 from Patent WO0155455.
ACCESSION AX207158
VERSION AX207158.1 GI:15394946
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1
AUTHORS Jin, S.
TITLE Resistance sequences and uses thereof
JOURNAL Patent: WO 0155455-A 5 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)
FEATURES
SOURCE 1. 1400
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

BASE COUNT 437 a 313 c 278 g 372 t
ORIGIN

Query Match 0.9%; Score 25; DB 6; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 TAACAGCTTAGTATCATGCCAC 46
|||||
Db 29 TAACAGCTTAGTATCATGCCAC 53

RESULT 47
LOCUS AC098444 244749 bp DNA linear HTG 13-NOV-2002
DEFINITION Rattus norvegicus clone CH230-2G15, *** SEQUENCING IN PROGRESS ***
ACCESSION AC098444
VERSION AC098444.4 GI:24941288
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 244749)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Albrooke, S.L., Amaralunga, H.C., Are, J.R., Ayala, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C., Davy-carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Horn, P., Howard, S., Huber, J., Hult, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,
Karlsson, E., Kelly, S., Khan, U., King, J., Korvah, J., Kovar, C.,
Kraus, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Loubege, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenko, S., Ogun, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Qilles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojudo, K., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 244749)
Worley, K.C.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23664578.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: TUBB
Center clone name: CH230-2G15
----- Summary Statistics

Assembly program: Phrap, version 0.990329
Consensus quality: 231469 bases at least Q40
Consensus quality: 234436 bases at least Q30
Consensus quality: 236097 bases at least Q20
Estimated insert size: 242749; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 242587: contig of 242587 bp in length
* 242588 242587: gap of unknown length
* 242688 244749: contigs of 2062 bp in length.
Location/Qualifiers

FEATURES

source

1. 244749
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-2G15"

misc_feature

1. 1633
/note="wgs end extension
clone_end:5p6"

misc_feature

/note="wgs end extension
clone_end:5p6"

misc_feature

complement(5662..6488)
/note="clone boundary
clone_end:5p6
site:
end sequence: BH287303"
complement(239764..240641)
/note="clone boundary
clone_end:t7
site:
end sequence: BH287302"

misc_feature

complement(239764..240641)
/note="clone boundary
clone_end:t7
site:
end sequence: BH287302"

BASE COUNT 78502 a 43428 c 44265 g 71430 t 7124 others
ORIGIN

Query Match 0.9%; Score 25; DB 2; Length 244749;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1791 GAGACCTTACTGCTGCTGGGATG 1815
|||||

Db 228396 GAGACCTTACTGCTGCTGGGATG 228372
|||||

RESULT 48
AC096525 250672 bp DNA linear HTG 10-MAY-2003
LOCUS Rattus norvegicus clone CH230-84N24, *** SEQUENCING IN PROGRESS
DEFINITION *** 3 unordered pieces.

ACCESSION

AC096525
AC096525.7 GI:30522022

VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 250672)
Munz, D., Marie, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alebrooks, S., Amin, A., Argüello, D.,
Ayalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Bernabed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, D.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draeper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lourenshewa, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinley, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L., Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokwelen, O., Okumu, G., Olarnpunagoon, A., Pal, S., Parks, K., Patel, S., Paul, H., Perez, A., Perez, L., Pflannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, B., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, D., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 250672)
Worley, K. C.

Direct Submission
Submitted (18-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 250672)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24959330.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CGBP
Center clone name: CH230-84N24
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 229622 bases at least Q40
Consensus quality: 232300 bases at least Q30

Consensus quality: 234233 bases at least Q20
Estimated insert size: 241730; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 237667: contig of 237667 bp in length
* 237668 237667: gap of unknown length
* 237768 247530: contig of 9763 bp in length
* 247531 247630: gap of unknown length
* 247631 250672: contig of 3042 bp in length.

FEATURES
source
Location/Qualifiers
1. 250672
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-84N24"
1. 2033
/note="wgs contig"

misc_feature
237768..238784
/note="wgs contig"

misc_feature
44292..43115 g 74893 t 15561 others

BASE COUNT 72811 a 44292 c 43115 g 74893 t 15561 others
ORIGIN

Query Match 0.9%; Score 25; DB 2; Length 250672;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 25; Conservative 0; Mismatches 0; Gaps 0;

Qy 1791 GAGACCTTACTGCTGGGATG 1815
Db 170733 GAGACCTTACTGCTGGGATG 170757

RESULT 49
AX591128 24 bp DNA linear PAT 27-JAN-2003
LOCUS AX591128
DEFINITION Sequence 1 from Patent WO02086157.
ACCESSION AX591128
VERSION AX591128.1 GI:27949643
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1. Latif, A., Cussenot, O. and Algarre-Genin, M.
AUTHORS Use of semaphorin 3a for diagnosing and treating cancer, especially
TITLE prostate cancer
JOURNAL Patent: WO 02086157-A 1 31-OCT-2002;
JOURNAL Urogene Societe anonyme (FR)
FEATURES
source Location/Qualifiers
1. 24
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="Amorce d'amplification du gene de la
semaphorine-3A (upper primer) choisie avec l'assistance des
programmes informatiques Oligo4 (Na tional BioSciences,
Plymouth, MN) et Primer express (Perkin-El mer Applied
BioSystems, Foster city, CA)"

BASE COUNT 8 a 7 c 4 g 5 t
ORIGIN
Query Match 0.9%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 24; Conservative 0; Mismatches 0; Gaps 0;

QY 1442 CCTATGACATGCCCCAATAGTG 1465
LOCUS |||||
Db 1 CCTATGACATGCCCCAATAGTG 24

RESULT 50
GGU02528 3263 bp mRNA linear VRT 01-NOV-1993
LOCUS gallus collapsin mRNA, complete cds.
DEFINITION U02528
ACCESSION U02528.1 GI:410078
VERSION
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 3263)
AUTHORS Luo, Y., Raible, D. and Raper, J.A.
TITLE Collapsin: a protein in brain that induces the collapse and
paralysis of neuronal growth cones
JOURNAL Cell 75 (2), 217-227 (1993)
MEDLINE 94006554
PubMed 8402908
REFERENCE 2 (bases 1 to 3263)
AUTHORS Luo, Y.
TITLE Direct Submission
SUBMITTED (13-OCT-1993) Yuling Luo, Department of Neuroscience,
University of Pennsylvania, School of Medicine, Philadelphia, PA
19104 USA

FEATURES
source
Location/Qualifiers
1..3263
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/sex="female"
/cissue_type="brain"
/clone_lib="chick brain cDNA library (Clontech Cat.
#CL1016A)"
/dev_stage="adult"
168..2486
/codon_start=1
/product="collapsin"
/protein_id="AAC59638.1"
/db_xref="GI:410078"

CDS
/translation="MGMRLGILLSGLVLAQVNCQHYKNNVPRKLSYKMLSSNN
IYNFGLASSSYHTFLDEERSRLVGAADHIFSNLVNIEYQKIVPVSRRDE
CEWAGKDIIRKCANFLKVLKTVNQTHLVACGAFHFMCTYEVSHPEDNFRMDS
HEWNGKSPYDPKLTJSLVDELSTADFMGRDPAIFRTLGHHPIRTEDHS
RWINDPRFSAHLIPESDDEDDKIYFFRENALDGEHKGATHARIGQICNDGEGH
RLVAKMTFLKRLICSVPGNGIDTHEDELQDVLMSKQKPKPIVGVGTTSSNI
FGSAVCTSMVDVRVRFGRPAHRDGPYQVPPYOGRIYPRRPGTCEPKTGPDST
KULPDEVITFARSHPMVNPVPIINSRPMITVDVQTOIVDVEDQDYDME
IGTDIGTVLKVVISPEKTEWHELEVLLEMTVREPTVISAKEISKQOOLVIGSAG
VSQLPHRCVDVYGAKCAECCLARDPYCAMDSSCSGRFPYAKRRRQINDKPLTH
CSDLGHNDPSCGLEEKIYIGVENSSTFLECSKPSQRAIVVWOKOMDDEKIKV
DDRMIRTEGGLIRSLQRDSDGIYFCHAVEHGFIOILLVITTEVIDTHLELLKKEE
DDADSKTKATNTSMTPSKITWRDKNQILNHNINMTMDFECQWKKRDKORORPAN
AQNVTNKGHLENNKGNRRNRHFERAPRSV"

BASE COUNT 1007 a 681 c 713 g 862 t
ORIGIN

Query Match 0.9%; Score 24; DB 5; Length 3263;
Best Local Similarity 100.0%; Pred No. 0.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 535 AGAATGCTAATTCATCAAGT 558
|||||
Db 503 AGAATGCTAATTCATCAAGT 526
|||||

RESULT 51

AC102985
LOCUS
DEFINITION Rattus norvegicus clone CH230-155c17, *** SEQUENCING IN PROGRESS
*** 3 unordered pieces.
ACCESSION AC102985.5 GI:30521923
VERSION
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 225005)
AUTHORS Muzny, D., Marie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alperocks, S., Amlin, A., Angiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., p'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, U., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dith, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorjis, E., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C.,
Kows, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Louissege, H., Lozardo, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Natir, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackemeah, O., Okumou, G., Olarunpasegun, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindecker, A., Popovic, D., Primbs, B., Pu, L., L.,
Punzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,
Stelmie, M., Strong, R., Sutton, A., Svatek, A., Taber, C., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wiczcyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 225005)
TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
TITLE
REFERENCE
AUTHORS
TITLE
Direct Submission
Submitted (26-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225005)
Rat Genome Sequencing Consortium.
Direct Submission

JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23664931.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHDG
Center clone name: CH230-155C17

----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 215727 bases at least Q40
Consensus quality: 218302 bases at least Q30
Consensus quality: 219785 bases at least Q20
Estimated insert size: 233913; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 218040: contig of 218040 bp in length
* 218041 218140: gap of unknown length
* 218141 222614: contig of 4474 bp in length
* 222615 222714: gap of unknown length
* 222715 225005: contig of 2291 bp in length.
Location/Qualifiers

FEATURES
SOURCE
1. 225005
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-155C17"

misc_feature

misc_feature

misc_feature

misc_feature

BASE COUNT 66589 a 44892 c 44710 g 64099 t 4715 others
ORIGIN

Query Match 0.9%; Score 24; DB 2; Length 225005;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1075 TTGCTCAGTGCAGGTCCAAATGG 1098
Db 127822 TTGCTCAGTGCAGGTCCAAATGG 127845

RESULT 52

AX591132 29 bp DNA linear PAT 27-JAN-2003
LOCUS Sequence 5 from Patent WO02086157.
DEFINITION AX591132
ACCESSION AX591132
VERSION AX591132.1 GI:27949647
KEYWORDS

ORGANISM

unidentified
unidentified
unclassified.

REFERENCE 1
AUTHORS Lactil A., Cussenot O. and Algate-Genin M.
TITLE Use of semaphorin 3a for diagnosing and treating cancer, especially prostate cancer
JOURNAL Patent: WO 02086157-A 5 31-OCT-2002;
trogene Societe anonyme (PR)

FEATURES

SOURCE

Location/Qualifiers
1. 29
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

BASE COUNT 6 a 6 c 9 g 8 t
ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

192 TCTCAGCATGCGCTGGTTAACT 214
Db 7 TCTCAGCATGCGCTGGTTAACT 29

Db

RESULT 53

D21099 1822 bp mRNA linear ROD 05-FEB-1999
LOCUS Mouse mRNA for STK-1 (serine/threonine kinase), complete cde.
DEFINITION D21099
ACCESSION D21099
VERSION D21099.1 GI:2182131
KEYWORDS

MSSTK1; serine/threonine kinase; STK-1; protein kinase.
MUS musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1822)
Nlwa,H., Abe,K., Kunisada,T. and Yamamura,K.
Cell-cycle-dependent expression of the STK-1 gene encoding a novel murine putative protein kinase

JOURNAL

Gene 169 (2), 197-201 (1996)
MEDLINE 8647446
PubMed 8647446

REFERENCE

2 (bases 1 to 1822)
Nlwa,H.
Direct Submission
Submitted (06-OCT-1993) Hitoshi Niwa, I. M. E. G. Kumamoto
University School of Medicine, Department of Developmental
Genetics; 4-24-1 Kuhonji, Kumamoto, Kumamoto 862, Japan
(Tel:096-344-2111(ex.5320), Fax:096-373-5321)
On Jun 8, 1997 this sequence version replaced gi:1322143.
Location/Qualifiers

COMMENT 1. 1822
FEATURES
SOURCE
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/tissue_type="testis"

```

gene      /dev_stage="adult"
          1. .1822
          /gene="MMSTK1"
CDS       184. .1221
          /gene="MMSTK1"
          /note="putative serine/threonine kinase"
          /product="STK-1"
          /codon_start=1
          /protein_id="BAA04658.1"
          /db_xref="GI:1322144"
          /translation="MAOKENAYPMYGSKTSGSGLNTLSORVIREKAPATSLALAVN
          SNSQSTAPAGQKIAENKSGSSTSGSOKQPTTINFEIGPLGKRGNYTLAREK
          KSRFVALKILEKSGLEKEGVHQLRELEIQAHLEKPNILQYNYFYQQRH YILLE
          YAPRGELAYLEKSRTEDEORTATIMEELSDALYCHKKKVLRHIDIKPENLLIGLQGE
          LKADGMSVHAASLRKTKMTGTLDTLPMEIEGRHNMVMDLMCGVLCYELMGNP
          PPESPSHSEYRIYAVDLPKPPSVSPSGQDLSIKLKINPNQRLPLAVAAHPWRA
          NSRRVLPSPAL"
          1794. .1799
          /gene="MMSTK1"
          1822
          /gene="MMSTK1"
polyA_site /gene="MMSTK1"
           435 a 497 c 449 g 441 t
BASE COUNT
ORIGIN
Query Match      0.8%; Score 23; DB 10; Length 1822;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCGCGCGTCCGGAGTA 79
    |||||
Db 22 GTGACCCGCGCGTCCGGAGTA 44

RESULT 54
AY030051 2325 bp mRNA linear VRT 01-MAY-2002
DEFINITION Xenopus laevis semaphorin 3A mRNA, complete cds.
ACCESSION AY030051
VERSION AY030051.1 GI:20384676
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2325)
Xenopus laevis
Tannahill, D., Nielsen, J. and Regan, A.G.
Xenopus Semaphorin 3A
Unpublished
2 (bases 1 to 2325)
Tannahill, D., Nielsen, J. and Regan, A.G.
Direct Submission
Submitted (11-APR-2001) Anatomy, University of Cambridge, Downing
Street, Cambridge CB2 3DY, UK
Location/Qualifiers
1. .2325
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
1. .2325
/tisue_type="head"
/note="class 3 secreted semaphorin; axonal guidance
protein"
/codon_start=1
/product="semaphorin 3A"
/protein_id="AAK38166.1"
/db_xref="GI:20384677"
/cis_xref="GI:20384677"
/translation="MGSLSLWIAFLFGLVSLITRLNCONVKNVPRILRLSYKEMVSN
LITFNGWNSSTYFVLDSEGRLLYVGAQKHIFSNLANVKEPQKILAMPVNIIRDE
CKMAGKDIYKECANFIVLKAKNHTHLVACGAFHPVCTYVDGHVEPDONVFRLED
SFEENGKRSKSPYDPLTASLILDELYSGTADMGDFRFAFRLTGNHPIRTQCHD
SRMLNDPRFICAYIVPESNDPDDKYVFFFRNATIDGHTGATHTARIGQCKNDPFG
HSLVKNWTTFLKARLCSVPGPNGIDTHDELQDVLMNSKDPNPVAVAVFTTSSN

```

```

IFKSAVCMYSLTDIRRVFLGPAHARDGPNTQWPFQGRVYPRPCTCPROTFGFDS
TKDLDEVIWFAHSHAPMYNVPFPINNPILIKTEVDYQFOIVDRVLEADGQYDM
FIGTDMGTILKVVSVKETWTDLEBTEWPEPEPAISAMESTOQOOLYIGSSV
GVSOILPRCDVYGKACCAECLADPYCAMGSSGSRFPAPKATRRPQDILRNGDPLT
HCSDDQHDDPHRQSLERKITYCHNSSTPIECSSKSRALVFPQKQNEKDEIK
VDERIKTEHGLTLTKRDSGLYYCAVHEGFQTLTKTLETIDHELDLHKE
DEGDSHKHKEPSNMSPTOKIWRDPMQLNHPNLTMDEFCQVWRDRKQRQRKN
GNVQVSNTKWHLQENKGRNRRWTFEPRARBSV"
BASE COUNT      744 a 444 c 517 g 620 t
ORIGIN
Query Match      0.8%; Score 23; DB 5; Length 2325;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 GCAAGGATCATATTTCATT 438
    |||||
Db 217 GCAAGGATCAATATTTCATT 239

RESULT 55
AC100365 51312 bp DNA linear HTG 22-NOV-2001
LOCUS
DEFINITION Mus musculus clone RP23-128120, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC100365
VERSION AC100365.1 GI:17047731
KEYWORDS HTG; HTGS PHASE0.
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathu; Muridae; Murinae; Mus.
1 (bases 1 to 51312)
Mus musculus, clone RP23-128120
Unpublished
2 (bases 1 to 51312)
Birken, B., Linton, L., Nusbbaum, C. and Lander, E.
Birken, B., Linton, L., Nusbbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Chospel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C.,
McCarthy, M., McKean, P., McKernan, K., McNeeters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zemlek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15170
Center clone name: 128_1_20

```

* NOTE: This record contains 65 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 687: contig of 687 bp in length
 688
 787: gap of 100 bp
 788
 1479: contig of 692 bp in length
 1480
 1579: gap of 100 bp
 1580
 2256: contig of 677 bp in length
 2257
 2356: gap of 100 bp
 2357
 3050: contig of 694 bp in length
 3051
 3150: gap of 100 bp
 3151
 3830: contig of 680 bp in length
 3831
 3930: gap of 100 bp
 3931
 4612: contig of 682 bp in length
 4613
 4712: gap of 100 bp
 4713
 5374: contig of 662 bp in length
 5375
 5474: gap of 100 bp
 5475
 6159: contig of 685 bp in length
 6160
 6259: gap of 100 bp
 6260
 6979: contig of 720 bp in length
 6980
 7079: gap of 100 bp
 7080
 7753: contig of 674 bp in length
 7754
 7853: gap of 100 bp
 7854
 8520: contig of 667 bp in length
 8521
 9307: contig of 687 bp in length
 9308
 9407: gap of 100 bp
 9408
 10103: contig of 636 bp in length
 10104
 10203: gap of 100 bp
 10204
 10912: contig of 703 bp in length
 10913
 11012: gap of 100 bp
 11013
 11688: contig of 676 bp in length
 11689
 11788: gap of 100 bp
 11789
 12490: contig of 702 bp in length
 12491
 12590: gap of 100 bp
 12591
 13283: contig of 693 bp in length
 13284
 13383: gap of 100 bp
 13384
 14049: contig of 666 bp in length
 14050
 14149: gap of 100 bp
 14150
 14837: contig of 688 bp in length
 14838
 14937: gap of 100 bp
 14938
 15643: contig of 706 bp in length
 15644
 15743: gap of 100 bp
 15744
 16442: contig of 699 bp in length
 16443
 16542: gap of 100 bp
 16543
 17220: contig of 678 bp in length
 17221
 17320: gap of 100 bp
 17321
 18009: contig of 689 bp in length
 18010
 18109: gap of 100 bp
 18110
 18814: contig of 705 bp in length
 18815
 18914: gap of 100 bp
 18915
 19602: contig of 688 bp in length
 19603
 19702: gap of 100 bp
 19703
 20390: contig of 688 bp in length
 20391
 20490: gap of 100 bp
 20491
 21173: contig of 683 bp in length
 21174
 21273: gap of 100 bp
 21274
 21945: contig of 672 bp in length
 21946
 22045: gap of 100 bp
 22046
 22742: contig of 697 bp in length
 22743
 22842: gap of 100 bp
 22843
 23541: contig of 699 bp in length
 23542
 23641: gap of 100 bp
 23642
 24340: contig of 699 bp in length
 24341
 24440: gap of 100 bp

FEATURES

SOURCE

24441
 25126
 25226
 25937: contig of 712 bp in length
 25938
 26037: gap of 100 bp
 26038
 26733
 26832: gap of 100 bp in length
 26833
 27518
 27617: gap of 100 bp in length
 27618
 28286: contig of 669 bp in length
 28287
 28386: gap of 100 bp
 28387
 29103: contig of 717 bp in length
 29104
 29203: gap of 100 bp
 29204
 29886: contig of 683 bp in length
 29887
 29986: gap of 100 bp
 30670: contig of 684 bp in length
 30671
 30770: gap of 100 bp
 30771
 31493: contig of 723 bp in length
 31494
 31593: gap of 100 bp
 31594
 32274
 32373: contig of 680 bp in length
 32374
 33078: contig of 705 bp in length
 33079
 33178: gap of 100 bp
 33179
 33877: contig of 699 bp in length
 33878
 33977: gap of 100 bp
 33978
 34669: contig of 692 bp in length
 34670
 34769: gap of 100 bp
 34770
 35414: contig of 645 bp in length
 35415
 35515
 35516: gap of 100 bp
 35517
 36222: contig of 708 bp in length
 36223
 36322: gap of 100 bp
 36323
 37007: contig of 685 bp in length
 37107: gap of 100 bp
 37108
 37830: contig of 723 bp in length
 37831
 37930: gap of 100 bp
 37931
 38632: contig of 702 bp in length
 38633
 38732: gap of 100 bp
 38733
 39392: contig of 660 bp in length
 39393
 39492: gap of 100 bp
 39493
 40155: contig of 663 bp in length
 40156
 40255: gap of 100 bp
 40256
 40963: contig of 708 bp in length
 40964
 41063: gap of 100 bp
 41064
 41763: contig of 700 bp in length
 41764
 41863: gap of 100 bp
 41864
 42581: contig of 718 bp in length
 42582
 42681: gap of 100 bp
 42682
 43385: contig of 704 bp in length
 43386
 43485: gap of 100 bp
 43486
 44201: contig of 716 bp in length
 44202
 44301: gap of 100 bp
 44302
 44978: contig of 677 bp in length
 44979
 45078: gap of 100 bp
 45079
 45771: contig of 693 bp in length
 45772
 45871: gap of 100 bp
 45872
 46574: contig of 703 bp in length
 46575
 46674: gap of 100 bp
 46675
 47371: contig of 697 bp in length
 47372
 47471: gap of 100 bp
 47472
 48164: contig of 693 bp in length
 48165
 48264: gap of 100 bp
 48265
 48951: contig of 687 bp in length
 48952
 49051: gap of 100 bp
 49052
 49722: contig of 671 bp in length
 49723
 49822: gap of 100 bp
 49823
 50517: contig of 695 bp in length
 50518
 51312: gap of 100 bp
 50618
 51312: contig of 695 bp in length.

Location/Qualifiers
 1..51312 "Mus musculus"
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-128120"

BASE COUNT 12962 a 8905 c 9030 g 13847 t 6568 others
 ORIGIN /clone_lib="RPCT-23 Female Mouse BAC"

Query Match 0.8%; Score 23; DB 2; Length 51312;
 Best Local Similarity 100.0%; Pred.No.1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 732 CATCCCTTTAATAGATGAGAA 754
 |||||
 DB 13056 CATCCCTTTAATAGATGAGAA 13078

RESULT 56
 AC120544/c 225372 bp DNA linear HTG 06-JUN-2003
 LOCUS Mus musculus clone RP23-299N6, *** SEQUENCING IN PROGRESS ***, 3
 DEFINITION Ordered pieces.
 AC120544
 AC120544.8 GI:31442497
 HTG: HTGS PHASE2; HTGS_FUILLTOP; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren, B., Nussbaum, C. and Lander, E.
 1 (bases 1 to 225372)
 Mus musculus, clone RP23-299N6
 Unpublished
 2 (bases 1 to 225372)

REFERENCE
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Bouhagbeler, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,
 Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fairo, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R.,
 Landers, T., Lehoczkay, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Meneses, L.,
 Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunhkhang, P., Pletre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE Submitted (08-MAY-2002) Whitehead Institute/MIT Center for Genome
 JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 225372)
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouhagbeler, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fairo, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Galham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meltrin, J., Meneses, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pletre, N.,
 Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schnapack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE Submitted (06-JUN-2003) Whitehead Institute/MIT Center for Genome
 JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE On Jun 6, 2003 this sequence version replaced gi:31339910.
 COMMENT All repeats were identified using RepeatMasker:
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 299_N6
 Center clone name: 299_N6

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 39864: contig of 39864 bp in length
 * 39865 39864: gap of 100 bp
 * 39965 74622: contig of 34658 bp in length
 * 74623 74722: gap of 100 bp
 * 74723 225372: contig of 150650 bp in length.
 Location/Qualifiers
 1..225372
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-299N6"
 /clone_lib="RPCT-23 Female Mouse BAC"
 BASE COUNT 67480 a 43750 c 43731 g 69944 t 467 others
 ORIGIN

Query Match 0.8%; Score 23; DB 2; Length 225372;
 Best Local Similarity 100.0%; Pred.No.1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 TGCTCAGTCCGCGTCCAAATGG 1098
 |||||
 DB 179450 TGCTCAGTCCGCGTCCAAATGG 179428

RESULT 57
 CGU28241
 LOCUS Gallus gallus 883 bp mRNA linear VRT 07-FEB-1996
 DEFINITION Gallus gallus collapsin-3 mRNA, partial cds.
 ACCESSION U28241
 VERSION U28241.1 GI:886810
 KEYWORDS
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 883)
 Luo, Y., Shepherd, I., Li, J., Renzi, M.J., Chang, S. and Raper, J.A.
 A family of molecules related to collapsin in the embryonic chick
 nervous system
 JOURNAL Neuron 14 (6), 1131-1140 (1995)
 MEDLINE 95329269
 PUBMED 7605628
 REFERENCE 2 (bases 1 to 883)
 AUTHORS Luo, Y.

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1995) Yuling Luo, Neuroscience, Univ.
Pennsylvania, 105 Johnson Pavilion, Philadelphia, PA 19104, USA

FEATURES
source
1.883
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
<1...>883
/codon_start=1
/product="collapsin-3"
/protein_id="AA86897.1"
/db_xref="GI:886811"
/translation="DDKIYFFEFKRLTNSGSKQIHSMARICPNDTGQORSLVINK
TTPFKARLVCSMDDEGTETFELEDEVLLETDPNPTLVYGIETSSSIFKSAVC
VYHLSDIOTVFNQSPRAHKGSGNHOLIIPYGRCTCGCAFTPMPTTKRPD
VTQIRNDPLATNYPPIHKKRPLIRITIDYKTKIANDRNADGRCHVIFLTGDDG
TVQKRVVLPINFSASGELILELEFQSNSPITTKISKQQLYSSSEGVTVPLH
RCRIYGTACACCLARDE"

BASE COUNT 268 a 185 c 195 g 235 t

ORIGIN

Query Match 0.8%; Score 22; DB 5; Length 883;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1778 TGTGGCTGGCCGACCTT 1799
|||||
Db 862 TGTGGCTGGCCGACCTT 883

RESULT 58
AX359932/c 5732 bp DNA linear PAT 13-FEB-2002

LOCUS AX359932
DEFINITION Sequence 6 from Patent WO0202783.
ACCESSION AX359932
VERSION AX359932.1 GI:18675571

KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS Bleck,G.T.
TITLE Expression vectors
JOURNAL Patent: WO 0202783-A 6 10-JAN-2002;
Gala Design, Inc. (US)
location/Qualifiers

FEATURES
source
1.5732
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic"

BASE COUNT 1390 a 1583 c 1425 g 1334 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 5732;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 5112 AATCTTTATTTATCGATGTT 5091

RESULT 59
AX382146/c 5732 bp DNA linear PAT 18-MAR-2002

LOCUS AX382146
DEFINITION Sequence 6 from Patent WO0202738.
ACCESSION AX382146
VERSION AX382146.1 GI:19576955

KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct

artificial sequences.

REFERENCE
1
AUTHORS Bremel,R.D., Miller,L.U., Bleck,G.T. and York,D.
TITLE Host cells containing multiple integrating vectors
JOURNAL Patent: WO 0202738-A 6 10-JAN-2002;
Gala Design, Inc. (US)
location/Qualifiers

FEATURES
source
1.5732
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 1390 a 1583 c 1425 g 1334 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 5732;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 5112 AATCTTTATTTATCGATGTT 5091

RESULT 60
A80476/c 7311 bp DNA circular PAT 21-JAN-2000

LOCUS A80476
DEFINITION Sequence 3 from Patent WO9943795.
ACCESSION A80476
VERSION A80476.1 GI:6731330

KEYWORDS
SOURCE
ORGANISM
unidentified
unidentified
unclassified.

REFERENCE
1 (bases 1 to 7311)
AUTHORS Benkel,B.F. and Falconer,M.M.
TITLE TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS
JOURNAL Patent: WO 9943795-A 3 02-SEP-1999;
BENKEL BERNHARD F (CA); FALCONER MARCIA M (CA)
location/Qualifiers

FEATURES
source
1.7311
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 1745 a 1935 c 1843 g 1788 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7311;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 4330 AATCTTTATTTATCGATGTT 4309

RESULT 61
BD136840/c 7311 bp DNA linear PAT 18-SEP-2002

LOCUS BD136840
DEFINITION Transsomatrics associated with gene transfer into mammary epithelial
cells.
ACCESSION BD136840
VERSION BD136840.1 GI:23231785
KEYWORDS JP 2002504365-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE
1 (bases 1 to 7311)
AUTHORS Gavora,J.S., Falconer,M.M., Nguyen,T.H. and Benkel,B.F.
TITLE Transsomatrics associated with gene transfer into mammary epithelial
cells
JOURNAL Patent: JP 2002504365-A 3 12-FEB-2002;
HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF
AGRICULTURE AND AGRI FOOD CANADA

COMMENT OS Unidentified
 PN JP 2002504365-A/3
 PD 12-FEB-2002
 PR 25-JUN-1998 JP 2000533535
 PR 24-FEB-1998 CA 2224108
 PI JAN S GAYORA, MARCIA M FALCONER, TUY H NGUYEN, BERNHARD F BENKEL
 PC A01K67/027, C12N9/64, C12N15/09, C12P21/00, C12N15/00 CC
 Strandedness: Double;
 Topology: Circular;
 CC Transonetics associated with gene transfer into mammary CC
 epithelial cells
 FH Key Location/Qualifiers
 FT source 1..7311
 FT Location/Qualifiers
 FT 1..7311
 FEATURES
 source 1..7311
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 1745 a 1935 c 1843 g 1788 t
 ORIGIN
 Query Match 0.8%; Score 22; DB 6; Length 7311;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AATCTTTATTTATCGATGTT 22
 ||||||||||||||||||
 4330 AATCTTTATTTATCGATGTT 4309
 RESULT 62
 AR302095/c 7352 bp DNA linear PAT 12-JUN-2003
 LOCUS
 DEFINITION Sequence 4 from patent US 6541197.
 ACCESSION AR302095
 VERSION AR302095.1 GI:31690129
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE
 1 (bases 1 to 7352)
 Link, C.J., Jr., Levy, J.P., Wang, S. and Seregina, T.
 Vehicles for stable transfer of green fluorescent protein gene and
 methods of use for same.
 Patent: US 6541197-A 4 01-APR-2003;
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..7352
 /organism="unknown"
 BASE COUNT 1735 a 1997 c 1877 g 1743 t
 ORIGIN
 Query Match 0.8%; Score 22; DB 6; Length 7352;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AATCTTTATTTATCGATGTT 22
 ||||||||||||||||||
 4371 AATCTTTATTTATCGATGTT 4350
 DB 4371 AATCTTTATTTATCGATGTT 4350
 RESULT 63
 AR302092/c 7353 bp DNA linear PAT 12-JUN-2003
 LOCUS
 DEFINITION Sequence 1 from patent US 6541197.
 ACCESSION AR302092
 VERSION AR302092.1 GI:31690126
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE
 1. (bases 1 to 7353)
 Link, C.J., Jr., Levy, J.P., Wang, S. and Seregina, T.

TITLE Vehicles for stable transfer of green fluorescent protein gene and
 methods of use for same
 JOURNAL Patent: US 6541197-A 1 01-APR-2003;
 FEATURES Location/Qualifiers
 source 1..7353
 /organism="unknown"
 BASE COUNT 1704 a 2062 c 1893 g 1694 t
 ORIGIN
 Query Match 0.8%; Score 22; DB 6; Length 7353;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AATCTTTATTTATCGATGTT 22
 ||||||||||||||||||
 4372 AATCTTTATTTATCGATGTT 4351
 DB 4372 AATCTTTATTTATCGATGTT 4351
 RESULT 64
 AX107929/c 7591 bp DNA linear PAT 30-APR-2001
 LOCUS
 DEFINITION Sequence 57 from Patent WO0125466.
 ACCESSION AX107929
 VERSION AX107929.1 GI:13923312
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE
 1
 Slingsby, J., Kingsman, S.M., Rohll, J.O. and Slade, A.O.
 Producer cell for the production of retroviral vectors
 Patent: WO 0125466-A 57 12-APR-2001;
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..7591
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 /note="Synthetic sequence"
 BASE COUNT 1752 a 2116 c 2007 g 1726 t
 ORIGIN
 Query Match 0.8%; Score 22; DB 6; Length 7591;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 AATCTTTATTTATCGATGTT 22
 ||||||||||||||||||
 3100 AATCTTTATTTATCGATGTT 3079
 DB 3100 AATCTTTATTTATCGATGTT 3079
 RESULT 65
 A80477/c 7885 bp DNA circular PAT 21-JAN-2000
 LOCUS
 DEFINITION Sequence 4 from Patent WO9943795.
 ACCESSION A80477
 VERSION A80477.1 GI:6731331
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE
 1 (bases 1 to 7885)
 Benkel, B.F. and Falconer, M.M.
 TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS
 TITLE PATENT: WO 9943795-A 4 02-SEP-1999;
 JOURNAL BENKEL, BERNHARD F (CA); FALCONER, MARCIA M (CA)
 FEATURES Location/Qualifiers
 source 1..7885
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 1876 a 2092 c 1996 g 1921 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7885;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 4904 AATCTTTATTTATCGATGTT 4883

RESULT 66
 BD136841/c 7885 bp DNA linear PAT 18-SEP-2002
 LOCUS Transsomatics associated with gene transfer into mammary epithelial
 DEFINITION cells.
 ACCESSION BD136841 GI:23231786
 VERSION JP 2002504365-A/4.
 KEYWORDS
 SOURCE unidentified
 ORGANISM

REFERENCE 1 (bases 1 to 7885)
 Gavora, J.S., Falconer, M.M., Nguyen, T.H. and Benkel, B.F.
 Transsomatics associated with gene transfer into mammary epithelial
 cells
 Patent: JP 2002504365-A 4 12-FEB-2002;
 HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF
 AGRICULTURE AND AGRI FOOD CANADA
 OS Unidentified
 PN JP 2002504365-A/4
 PD 12-FEB-2002
 PF 25-JUN-1998 JP 2000533535
 PR 24-FEB-1998 CA 2224108
 PI JAN S GAVORA, MARCIA M FALCONER, THUY H NGUYEN, BERNHARD F BENKEL,
 PC A01K67/027,C12N9/64,C12N15/09,C12P21/00,C12N15/00 CC
 Strandedness: Double;
 CC Topology: Circular;
 CC Transsomatics associated with gene transfer into mammary CC
 epithelial cells
 FH Key location/Qualifiers
 FT source 1..7885
 /organism='Unidentified'.
 Location/Qualifiers
 1..7885
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 1876 a 2092 c 1996 g 1921 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7885;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 4904 AATCTTTATTTATCGATGTT 4883

RESULT 67
 AX359933/c 9183 bp DNA linear PAT 13-FEB-2002
 LOCUS Sequence 7 from Patent WO0202783.
 DEFINITION
 ACCESSION AX359933
 VERSION AX359933.1 GI:18675572.
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Black, G.T.
 TITLE Expression vectors
 JOURNAL Patent: WO 0202783-A 7 10-JAN-2002;

Query Match 0.8%; Score 22; DB 6; Length 9183;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES Gala Design, Inc. (US)
 source 1..9183
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 /note="Synthetic"

Query Match 0.8%; Score 22; DB 6; Length 9183;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 8563 AATCTTTATTTATCGATGTT 8542

RESULT 68
 AX382147/c 9183 bp DNA linear PAT 18-MAR-2002
 LOCUS Sequence 7 from Patent WO0202738.
 DEFINITION
 ACCESSION AX382147
 VERSION AX382147.1 GI:19576956
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Bremer, R.D., Miller, L.U., Black, G.T. and York, D.
 Host cells containing multiple integrating vectors
 Patent: WO 0202738-A 7 10-JAN-2002;
 JOURNAL Gala Design, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..9183
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

BASE COUNT 2133 a 2498 c 2399 g 2153 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 9183;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 8563 AATCTTTATTTATCGATGTT 8542

RESULT 69
 AC006525/c 90604 bp DNA linear PRI 19-APR-2002
 LOCUS Homo sapiens chromosome 5 clone RP1-4017, complete sequence.
 DEFINITION
 ACCESSION AC006525
 VERSION AC006525.1 GI:4225899
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 90604)
 Kimmery, W., Bondoc, M., Cheng, J., Connolly, K.S., Kadner, K.,
 Mignel, T., Miller, C., Pitlick, S., Pollard, M., Rojeski, H.,
 Subramanian, S., Wheeland, A. and Martin, C.H.
 Sequencing of human chromosome 5
 Unpublished
 REFERENCE 2 (bases 1 to 90604)
 Rieke, D.O.
 Large Scale Sequence Analysis and Annotation with the Sequence
 Comparison Analysis (SCAN) System
 JOURNAL Unpublished

REFERENCE 3 (bases 1 to 90604)
AUTHORS Kimerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Kadner,K.,
Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojsek,I.H.,
Subramanian,S., Wheeland,A. and Martin,C.H.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1999) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
REFERENCE 4 (bases 1 to 90604)
AUTHORS Mundt,M.O.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2002) Biosciences, Joint Genome Institute/ Los
Alamos National Laboratory, MS M888, Los Alamos, New Mexico 87545,
U.S.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source
1. 90604
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP1-4017"
147. .264
/rpt_family="MIR"
repeat_region
871. .1157
/rpt_family="Alu"
repeat_region
complement(2190. .8204)
/rpt_family="L1"
repeat_region
complement(6746. .6964)
/rpt_family="MER25"
8436. .8715
/rpt_family="Alu"
repeat_region
9507. .9715
/rpt_family="MLT"
repeat_region
complement(11635. .11800)
/rpt_family="L1"
repeat_region
18197. .18493
/rpt_family="Alu"
repeat_region
complement(18701. .19051)
/rpt_family="THE1"
repeat_region
complement(19056. .19622)
/rpt_family="THE"
repeat_region
complement(19073. .20447)
/rpt_family="MSTAR"
repeat_region
complement(19655. .20251)
/rpt_family="THR"
misc_feature
complement(20961. .21052)
/note="GRAIL 2 excellent exon, frame 0"
22417. .22785
/rpt_family="THE1"
repeat_region
complement(24301. .25358)
/rpt_family="L1"
misc_feature
complement(26046. .26221)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region
complement(26626. .26830)
/rpt_family="L1"
repeat_region
complement(30565. .30815)
/rpt_family="MER44C"
repeat_region
complement(32234. .33083)
/rpt_family="L1"
repeat_region
complement(33304. .35878)
/rpt_family="L1"
repeat_region
complement(35886. .36070)
/rpt_family="Alu"
repeat_region
36037. .36270
/rpt_family="Alu"
repeat_region
complement(36268. .42451)
/rpt_family="L1"
repeat_region
complement(40878. .41365)
/rpt_family="MER25"
repeat_region
complement(42728. .42898)
/rpt_family="Alu"

repeat_region complement(42916. .43259)
/rpt_family="Alu"
repeat_region 44045. .44230
/rpt_family="MER3"
repeat_region 44096. .44142
/rpt_family="MER33"
repeat_region complement(45460. .45968)
/rpt_family="L1"
repeat_region complement(52284. .52407)
/rpt_family="MIR"
repeat_region complement(52989. .53321)
/rpt_family="Alu"
repeat_region complement(53942. .54100)
/rpt_family="L1"
repeat_region 55602. .55886
/rpt_family="Alu"
repeat_region complement(56037. .56502)
/rpt_family="L1"
repeat_region 59874. .60193
/rpt_family="Alu"
repeat_region 60477. .60752
/rpt_family="Alu"
repeat_region 60820. .61300
/rpt_family="L1"
repeat_region 61348. .61619
/rpt_family="L1"
repeat_region 62747. .63044
/rpt_family="Alu"
misc_feature complement(63623. .63754)
/note="GRAIL 2 excellent exon, frame 0"
64221. .64406
/rpt_family="MER5"
repeat_region complement(64263. .64381)
/rpt_family="MER5"
repeat_region complement(66455. .66792)
/rpt_family="MER"
repeat_region complement(67054. .67310)
/rpt_family="Alu"
repeat_region complement(67511. .69174)
/rpt_family="L1"
repeat_region complement(69265. .69379)
/rpt_family="MIR"
repeat_region 70504. .73086
/rpt_family="L1"
repeat_region 71805. .72106
/rpt_family="Alu"
misc_feature 72865. .72937
/note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(73156. .74083)
/rpt_family="L1"
misc_feature 76196. .76310
/note="GRAIL 2 excellent exon, frame 2"
76342. .76625
/rpt_family="Alu"
repeat_region complement(77911. .78203)
/rpt_family="THE1"
repeat_region complement(78204. .79756)
/rpt_family="MSTAR"
repeat_region complement(79761. .80107)
/rpt_family="THE1"
repeat_region complement(83286. .83483)
/rpt_family="L1"
repeat_region complement(84022. .84091)
/rpt_family="L1"
repeat_region complement(84657. .84976)
/rpt_family="L1"
repeat_region 85589. .85681
/rpt_family="LTR"
repeat_region complement(86063. .86337)
/rpt_family="Alu"
repeat_region 86485. .86693
/rpt_family="L1"
repeat_region complement(87822. .90604)

```

BASE COUNT      25086 a 16222 c 16894 g 32402 t
ORIGIN
Query Match      0.8%; Score 22; DB 9; Length 90604;
Beet Local Similarity 100.0%; Pred. No. 4.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      287 CCAGGCTGAATATTCCTACA 308
DB      10922 CCAGGCTGAATATTCCTACA 10901

RESULT 70
AC112426.0
WPCOMMENT
Sequence split into 4 fragments LOCUS AC112426 Accession AC112426
Fragment Name      Begin      End
AC112426_0          1      110000
AC112426_1        100001      210000
AC112426_2        200001      310000
AC112426_3        300001      374204
LOCUS      AC112426      374204 bp      DNA      linear      HTG 08-OCT-2002
DEFINITION      Rattus norvegicus clone CH230-312K14, *** SEQUENCING IN PROGRESS
ACCESSION      AC112426
VERSION      AC112426.4 GI:23270041
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 374204)
Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S,, Adams,C,, Alder,J.,
Allen,C,, Allen,H,, Alsbrooks,S,, Amin,A,, Anguiano,D,,
Ayala-Lepech,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F.,
Biswal,K,, Blair,J., Blankenburg,K,, Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,J., Carter,K., Cavazos,I., Cesari,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Claveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Drapper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frieser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guayra,W.,
Guenarte,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpacky,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kovis,C., Kraft,C.L., Lebow,H., Lavan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louisgeed,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Matherly,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,G., Olarpunsaagoon,A., Pal,S., Parks,K.,
Nwokilemeh,O., Okwomou,G., Olarpunsaagoon,A., Pal,S., Parks,K.,
Palpatrak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L.,
Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojase,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sleson,I., Sitter,C.D., Smajs,D.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21743438.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atlas/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
genome sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRFB
Center clone name: CH230-312K14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 240799 bases at least Q40
Consensus quality: 246749 bases at least Q30
Consensus quality: 250548 bases at least Q20
Estimated insert size: 345761; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone
* NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 222049: contig of 222049 bp in length
222050 222149: gap of unknown length
222150 236912: contig of 14763 bp in length
236913 237012: gap of unknown length
237013 301105: contig of 64093 bp in length
301106 301205: gap of unknown length
301206 305403: contig of 4198 bp in length
305404 305503: gap of unknown length
305504 352310: contig of 46807 bp in length

```

```

* 352311 352410: gap of unknown length
* 352411 353860: contig of 1450 bp in length
* 353861 353960: gap of unknown length
* 353961 355587: contig of 1627 bp in length
* 355588 355687: gap of unknown length
* 355688 357943: contig of 2256 bp in length
* 357944 358043: gap of unknown length
* 358044 360096: contig of 2053 bp in length
* 360097 360196: gap of unknown length
* 360197 361828: contig of 1632 bp in length
* 361829 361928: gap of unknown length
* 361929 364193: contig of 2265 bp in length
* 364194 364293: gap of unknown length
* 364294 374204: contig of 9911 bp in length.

```

```

FEATURES
    source
        1..374204
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-312K14"
            /107683..110071
            /note="wgs contig"
            /220130..222049
            /note="wgs contig"
            /301206..303668
            /note="wgs contig"
            /303719..305403
            /note="wgs contig"

```

```

BASE COUNT  72927 a 53729 c 54751 g 71203 t 121594 others
ORIGIN

```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2248 AGATGCTTAAGACCAAGAA 2269
Db 91867 AGATGCTTAAGACCAAGAA 91888

```

```

RESULT 71
AC112426_2
WCOMMENT
Sequence split into 4 fragments LOCUS AC112426 Accession AC112426
Fragment Name Begin End
AC112426_0 1 110000
AC112426_1 100001 210000
AC112426_2 200001 310000
AC112426_3 300001 374204
Continuation (3 of 4) of AC112426 from base 200001 (AC112426 Rattus norvegicus clone CH2

```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2248 AGATGCTTAAGACCAAGAA 2269
Db 37340 AGATGCTTAAGACCAAGAA 37361

```

```

RESULT 72
AC129994 150638 bp DNA linear HTG 15-NOV-2002
LOCUS AC129994
DEFINITION Rattus norvegicus clone CH230-510F7, WORKING DRAFT SEQUENCE, 4
ACCESSION AC129994
VERSION AC129994.3 GI:25007674
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE AUTHORS

1 (bases 1 to 150638)
Muzny,D.,Marie, Metker,M.,Lee, Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angilano,D.,
Anyalebech,V., Ayogbi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cessari,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., p.Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Ande,C., Dederich,D.,
Delgado,O., Denson,S., Detamo,C., Ding,Y., Dint,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durkin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hawes,A., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jolivet,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louissegd,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Mundaasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokediemo,O., Okumu,G., Olampunagoo,A., Pal,S., Parks,K.,
Pascernak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
Plopper,F., Poindester,A., Popovic,D., Primis,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorrell,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vezar,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wiczcyk,R., Woodson,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Unpublished
2 (bases 1 to 150638)
Worley,K.C.
Direct Submission
Submitted (08-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23196240.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atl/) . Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: XAGT
Center clone name: CH230-510F7

Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 145002 bases at least Q40
Consensus quality: 146043 bases at least Q30
Consensus quality: 146587 bases at least Q20
Estimated insert size: 148398; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 128673: contig of 128673 bp in length
* 128674 128773: gap of unknown length
* 128774 147523: contig of 18750 bp in length
* 147524 147623: gap of unknown length
* 147624 148701: contig of 1078 bp in length
* 148702 148801: gap of unknown length
* 148802 150638: contig of 1837 bp in length.

FEATURES

source

1. 150638

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-510F7"

misc_feature

2017. 105883

/note="clone boundary"

clone_end:sp6

site:

end_sequence:B2117461"

6382. 7103

/note="clone boundary"

clone_end:t7

site:

end_sequence:B2117458"

128774. 131862

/note="wgs_end_extension"

clone_end:t7"

BASE COUNT 38301 a 36249 c 36133 g 36407 t 3548 others

ORIGIN

Query Match 0.8%; Score 22; DB 2; Length 150638;

Best Local Similarity 100.0%; Pred. No. 4.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GY 2461 GAAGCTGAAACAGAGACC 2482

Db 40496 GAAGCTGAAACAGAGACC 40517

RESULT 73

LOCUS

AC122712 166965 bp DNA linear HTG 25-MAY-2002

DEFINITION

Homo sapiens chromosome 5 clone RP11-35J11, WORKING DRAFT SEQUENCE,
19 unordered pieces.

ACCESSION

AC122712
AC122712.1 GI:21206275
HTG; HTGS_PHAESI; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE
Homo sapiens (human)

REFERENCE

1 (bases 1 to 166965)

2 (bases 1 to 166965)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

2 (bases 1 to 166965)

DOE Joint Genome Institute.

Direct Submission

Submitted (25-MAY-2002) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

COMMENT

Project Information
Center Project Name: 420032
Center clone name: RP11-35J11

Summary Statistics

Consensus quality: 149343 bases at least Q40
Consensus quality: 156729 bases at least Q30
Consensus quality: 160416 bases at least Q20

Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 165165; sum-of-contigs estimation

Quality coverage: 12.41 in Q20 bases; agarose-fp estimation
Quality coverage: 13.14 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1095: contig of 1095 bp in length
* 1096 1195: gap of unknown length
* 1196 2331: contig of 1136 bp in length
* 2332 2431: gap of unknown length
* 2432 3885: contig of 1454 bp in length
* 3886 3985: gap of unknown length
* 3986 5474: contig of 1489 bp in length
* 5475 5574: gap of unknown length
* 5575 6785: contig of 1211 bp in length
* 6786 6885: gap of unknown length
* 6886 8468: contig of 1583 bp in length
* 8469 8568: gap of unknown length
* 8569 11081: contig of 2513 bp in length
* 11082 11181: gap of unknown length
* 11182 14532: contig of 3351 bp in length
* 14533 14632: gap of unknown length
* 14634 18334: contig of 3702 bp in length
* 18335 18434: gap of unknown length
* 18435 28070: contig of 9636 bp in length
* 28071 28170: gap of unknown length
* 28171 34825: contig of 6655 bp in length
* 34826 34925: gap of unknown length
* 34926 43325: contig of 8400 bp in length
* 43326 43425: gap of unknown length
* 43426 56689: contig of 13264 bp in length
* 56690 56789: gap of unknown length
* 56790 70875: contig of 14086 bp in length
* 70876 70975: gap of unknown length
* 70976 70976: gap of unknown length
* 70977 86880: contig of 15905 bp in length
* 86881 86980: gap of unknown length
* 86981 106763: contig of 19783 bp in length
* 106764 106863: gap of unknown length

```

FEATURES
  source
    1.166965
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="5"
    /clone="RP11-35J11"
    /clone_1ib="RP11 human BAC library 11"
    /clone_2ib="31902 c 31191 g 53060 t 1807 others"

BASE COUNT
  49005 a 31902 c 31191 g 53060 t 1807 others

ORIGIN

Query Match
  Best Local Similarity 100.0%; Pred No. 4.4;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0.8%; Score 22; DB:2; Length 166965;

287 CCAGGCTGAATTATCTTACA 308
|||||
121337 CCAGGCTGAATTATCTTACA 121358

RESULT 74
AC098305/c
LOCUS
  AC098305
DEFINITION
  AC098305
ACCESSION
  AC098305
VERSION
  AC098305.6
KEYWORDS
  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE
  Rattus norvegicus (Norway rat)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
  1 (bases 1 to 216124)
  Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
  Alsprouk,S.L., Amarantunge,H.C., Are,J.R., Ayala,M., Banks,T.,
  Barbara,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
  Bouck,U., Bowles,S., Brileva,M., Brown,B., Brown,M., Bryant,N.P.,
  Burch,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
  Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
  Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
  Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
  Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
  Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
  Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
  Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
  Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Fratutz,P.,
  Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
  Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
  Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
  Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
  Homel,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
  Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
  Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
  Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
  Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Lousseng,H.,
  Lozada,R.U., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
  Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,B.,
  Masey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
  Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
  Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
  Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G.,
  Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
  Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
  Rivers,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savery,G.,
  Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
  Soedergen,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
  Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

```

```

TITLE
  JOURNAL
  Unpublished
  2 (bases 1 to 216124)
  Worley,K.C.
  Direct Submission
  Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 216124)
  Worley,K.C.
  Direct Submission
  Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA

REFERENCE
  Tansley,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N., Thomas,S.,
  Uemami,K., Vaequez,L., Vera,V., Villalon,D., Vinsom,R., Wang,Q.,
  Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
  Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
  Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
  Weinstein,G. and Gibbs,R.
  Direct Submission
  Unpublished
  2 (bases 1 to 216124)
  Worley,K.C.
  Direct Submission
  Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 216124)
  Worley,K.C.
  Direct Submission
  Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA

COMMENT
  The sequence in this assembly is a combination of BAC based reads
  and whole genome shotgun sequencing reads assembled using Atlas
  (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
  in the feature table below represents a scaffold in the Atlas
  assembly (a 'contig-scaffold'). Within each contig-scaffold,
  individual sequence contigs are oriented and separated
  by sized gaps filled with Ns to the estimated size. The sequence
  may extend beyond the ends of the clone and there may be sequence
  contigs within a contig-scaffold that consist entirely of whole
  genome shotgun sequence reads. Both end sequences and whole genome
  shotgun sequence only contigs will be indicated in the feature
  table.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: CH230-IN4
Center clone name: CH230-IN4
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329
Consensus quality: 132680 bases at least Q40
Consensus quality: 138597 bases at least Q30
Consensus quality: 142739 bases at least Q20
Estimated insert size: 145045; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length -----
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
  1
  67820: contig of 67820 bp in length
  67821
  67920: gap of unknown length
  67921
  71048: contig of 3128 bp in length
  71049
  71148: gap of unknown length
  71149
  130978: contig of 59830 bp in length
  130979
  131078: gap of unknown length
  131079
  131079: contig of 66392 bp in length
  131079
  131079: gap of unknown length
  131079
  131079: contig of 5646 bp in length
  131079
  205216: contig of 5646 bp in length
  205217
  205316: gap of unknown length
  205317
  206488: contig of 1532 bp in length
  206489
  206489: gap of unknown length
  206490
  206490: contig of 1470 bp in length
  206491
  208518: gap of unknown length
  208519

```

FEATURES

source

1. 216124
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1N4"

BASE COUNT 41454 a 31758 c 30503 g 41391 t 71018 others

ORIGIN

Query Match 0.8%; Score 22; DB 2; Length 216124;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2248 AGATGGCTTAAGACCAAGAA 2269
DB 189397 AGATGGCTTAAGACCAAGAA 189376

RESULT 75
AC106106/c

LOCUS AC106106 222086 bp DNA linear HTG 10-MAY-2003

DEFINITION Rattus norvegicus clone CH230-14317, WORKING DRAFT SEQUENCE.

ACCESSION AC106106

VERSION AC106106.5 GI:30521604

KEYWORDS HTG_PHASE2; HTGS_DRAFT; HTGS_FULFILLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus; Rattus

REFERENCE 1 (bases 1 to 222086)
Munzy, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, D., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lounsbury, L., Louised, H., Lozado, R. D., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangun, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mauney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokediemeh, O., Okwomou, G., Olarnpungoon, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

SAUNDERS, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Sivartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Soderstrom, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 222086)
Worley, K. C.

Direct Submission
Submitted (12-UN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 222086)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:25094802.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GHKJ
Center clone name: CH230-14317
----- Summary Statistics
Assembly program: Atlas 3.0j
Consensus quality: 21592 bases at least Q40
Consensus quality: 216173 bases at least Q30
Consensus quality: 217783 bases at least Q20
Estimated insert size: 225854; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 222086: contig of 222086 bp in length.
Location/Qualifiers
1. 222086
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-14317"


```
misc_feature 1..3026
BASE COUNT 55222 a 54777 c 53885 g 55298 t 2904 others
ORIGIN
Query Match 0.8%; Score 22; DB 2; Length 222086;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2461 GAAGGTAGAAACAGGAGACC 2482
Db 26127 GAAGGTAGAAACAGGAGACC 26106
```

Search completed: July 31, 2003, 22:28:36
Job time : 9595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:21:21 ; Search time 5030 Seconds
(without alignments)
13089.627 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatctttatctatcatgacg.....agcttttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_dhg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	ID	Description
1	723	26.7	723	12	BI870437
2	526	19.4	679	12	BM711125
3	418	15.4	702	12	BM723807
4	408	15.1	510	14	CD216784

5	391	14.4	889	10	BF667677
6	387	14.3	497	14	CA3955261
7	382	14.1	392	13	CA893116
8	324	12.0	721	12	HS084323
9	315	11.6	670	12	BF667627
10	286	10.6	839	10	BF667143
11	284	10.5	747	12	BQ018643
12	274	10.1	884	10	BF700780
13	265	9.8	535	2	HS084322
14	260	9.6	692	12	HS084322
15	218	8.0	355	10	BF742853
16	213	7.9	231	14	T29595
17	203	7.5	328	9	A1671991
18	196	7.2	210	14	R19784
19	194	7.2	914	10	AL529332
20	135	5.0	325	10	BE552070
21	120	4.8	430	9	AA424316
22	104	3.8	896	13	B0186364
23	101	3.7	570	10	BE033028
24	68	2.5	1005	12	BM801410
25	67	2.5	677	29	AG157495
26	59	2.2	604	12	BI183160
27	57	2.1	598	13	EX282800
28	56	2.1	576	28	AQ533647
29	52	1.9	559	14	CA872840
30	51	1.9	394	9	AA716143
31	50	1.8	773	28	BH265538
32	49	1.8	124	10	BG315503
33	47	1.7	371	13	BY021039
34	47	1.7	401	13	BY018262
35	47	1.7	450	13	BY261117
36	47	1.7	470	12	BI465498
37	47	1.7	639	10	BB655104
38	47	1.7	653	10	BB634470
39	47	1.7	655	10	BB667377
40	47	1.7	657	10	BB658393
41	47	1.7	677	14	BY734980
42	47	1.7	731	13	BU708579
43	47	1.7	787	14	CB518391
44	47	1.7	2596	11	AK053115
45	41	1.5	318	13	BY303092
46	41	1.5	807	14	CB595903
47	41	1.5	900	14	CB196368
48	41	1.5	917	13	BO884924
49	38	1.4	394	14	CB776561
50	37	1.4	657	13	BQ779621
51	37	1.4	658	13	BU758296
52	37	1.4	660	10	BB612039
53	35	1.3	699	14	CA327319
54	35	1.3	773	13	BU515916
55	35	1.3	775	13	BU703431
56	34	1.3	374	9	AA861394
57	30	1.1	600	12	BI989239
58	30	1.1	662	10	BB628362
59	30	1.1	690	14	W98303
60	30	1.1	712	9	AI894009
61	29	1.1	291	10	BB377317
62	29	1.1	292	10	BB366496
63	29	1.1	729	14	CD349263
64	27	1.0	446	10	BB859723
65	27	1.0	539	9	AL918018
66	25	0.9	685	12	BG926854
67	24	0.9	555	14	CD203422
68	24	0.9	598	14	CD203073
69	24	0.9	614	14	CA502146
70	24	0.9	792	12	BG781873
71	24	0.9	825	13	BU255423
72	24	0.9	851	13	BU326664
73	24	0.9	897	13	BU267223
74	23	0.8	184	9	AM080828
75	23	0.8	401	12	BG898214
76	23	0.8	591	14	CA779226
77	23	0.8	608	13	BQ907117

BE667677	602121773
CA3955261	CS62C03.Y
BX283116	BX283116
BX510254	Homo sapi
BM670627	UI-E-DX1-
BF667143	602121928
BQ018643	UI-H-DH1-
BF700780	602128585
Bx491494	Homo sapi
BQ447821	UI-H-EU1-
BF742853	IL2-BT073
T29595	EST86074.Hu
A1671991	wb72C04.X
R19784	YG28C12.r1
AL529332	AL529332
BE552070	hy03d11.X
AA424316	zv90D08.Y
B0186364	AGENCOURT
BE033028	133242.MA
BM801410	AGENCOURT
AG157495	Par trogl
BI183160	UNI-P-FN-
EX282800	BX282800
AQ533647	RPCI-11-3
CA872840	K0920H08-
AA716143	zg60D02.S
BH265538	CH230-43J
BG315503	PO3.0.131
BY021039	BY021039
BY018262	BY018262
BY261117	BY261117
BI465498	ie18d11.Y
BB655104	BB655104
BB634470	BB634470
BB667377	BB667377
BB658393	BB658393
BY734980	BY734980
BU708579	UI-M-F10-
CB518391	UI-M-GH0-
AK053115	Mus muscu
BY303092	BY303092
CB595903	AGENCOURT
CB196368	AGENCOURT
BO884924	AGENCOURT
CB776561	AMGNIC.S
BQ779621	UI-R-FP0-
BU758296	UI-R-FP0-
BB612039	BB612039
CA327319	UI-M-FY0-
BU515916	AGENCOURT
BU703431	UI-M-FP0-
AA861394	ak37C08.S
BI989239	4032-91.M
BB628362	BB628362
W98303	mg13h01.r1
AI894009	mg13h01.Y
BB377317	BB377317
BB366496	BB366496
CD349263	UI-M-FY0-
BB859723	BB859723
AL918018	AL918018
BG926854	HNC20-1-B
CD203422	LG_AM1_08
CD203073	LG_AM1_02
CA502146	WHR4043.C
BG781873	SEADMC001
BU255423	603414282
BU326664	603493391
BU267223	603815106
AM080828	XC38A05.X
BG898214	H0A38-1-B
CA779226	MDP384_8
BQ907117	N003D03.O

78	23	0.8	625	12	BG897406	BG897406 HOA13-1-F
79	23	0.8	638	12	BG926247	BG926247 HNC50-1-G
80	23	0.8	730	12	BG975571	BG975571 602845241
81	23	0.8	822	12	BF615577	BF615577 603305839
82	22	0.8	180	10	BF984015	BF984015 602306942
83	22	0.8	213	10	BE609913	BE609913 sq46508.Y
84	22	0.8	225	10	BB368206	BB368206 BB368206
85	22	0.8	330	9	AU036240	AU036240 AU036240
86	22	0.8	366	13	BG969884	BG969884 LL21n3001
87	22	0.8	408	12	BG897335	BG897335 HOA12-1-F
88	22	0.8	409	12	BG926168	BG926168 HNC2-1-D3
89	22	0.8	410	12	BG924214	BG924214 HNC25-1-G
90	22	0.8	415	10	BG025774	BG025774 602274765
91	22	0.8	418	12	BG897329	BG897329 HOA12-1-E
92	22	0.8	418	12	BG898828	BG898828 HOA57-1-C
93	22	0.8	435	12	BI813570	BI813570 KO14F09.O
94	22	0.8	441	14	CA871124	CA871124 KO908A03-C
95	22	0.8	447	12	BG901133	BG901133 HOA54-1-B
96	22	0.8	452	12	BG926300	BG926300 HNC3-1-E8
97	22	0.8	464	14	CA779986	CA779986 MFL384.3
98	22	0.8	474	12	BG927549	BG927549 HNC43-1-B
99	22	0.8	492	10	BF162689	BF162689 HNC43-1-B
100	22	0.8	493	14	CB725110	CB725110 AMGNNUC.N

ALIGNMENTS

RESULT 1
BI870437
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI870437 723 bp mRNA linear EST 11-OCT-2001
603395690F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405519 5',
mRNA sequence.
BI870437
BI870437.1 GI:16044110
EST.
Homo sapiens (human)
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 723)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM2034 row: C column: 24
High quality sequence stop: 723.
Location/Qualifiers
1: 723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405519"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_90"
/note="Organ: Liver; Vector: pCMV-SORTe; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."

FEATURES
source
BASE COUNT
ORIGIN
Query Match %

26.7%; Score 723; DB 12; Length 723;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	454	CAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGATGATGCAAGG	513
DB	1	CAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGATGCAAGG	60
QY	514	GCGTGAAGAAGACATCTTGAAGAATGTGCTAATTTTCATCAAGTAACTTAAGGCAATATA	573
DB	61	GCGTGAAGAAGACATCTTGAAGAATGTGCTAATTTTCATCAAGTAACTTAAGGCAATATA	120
QY	574	TGAGACTCACTTGTAGCGCTGTGAAACGGGGCTTTTCATCAATTTGCACTTAATTGA	633
DB	121	TGAGACTCACTTGTAGCGCTGTGAAACGGGGCTTTTCATCAATTTGCACTTAATTGA	180
QY	634	AATTTGACATCATCTTGAAGGAGCAATATTTTAAGCTGAGAACTCAATTTGAAAAGG	693
DB	181	AATTTGACATCATCTTGAAGGAGCAATATTTTAAGCTGAGAACTCAATTTGAAAAGG	240
QY	694	CCGTGGGAAGAGTCCATATGACCTTAAGCTGACAGCATCCCTTTAATAGATGAGA	753
DB	241	CCGTGGGAAGAGTCCATATGACCTTAAGCTGACAGCATCCCTTTAATAGATGAGA	300
QY	754	ATTATCTCTGAACTGACGTGATTTTATGGGGGAGACTTTGCTATCTTCCGAATCT	813
DB	301	ATTATCTCTGAACTGACGTGATTTTATGGGGGAGACTTTGCTATCTTCCGAATCT	360
QY	814	TGGGACACCAACCAATAGAGAGAGAGATGATTCAGGTGGCTTAATGCCAA	873
DB	361	TGGGACACCAACCAATAGAGAGAGATGATTCAGGTGGCTTAATGCCAA	420
QY	874	GTTCAATTAGTGGCCACCTCATCTCAGAGATGACATCTTGAAGATGACAAAGTATCT	933
DB	421	GTTCAATTAGTGGCCACCTCATCTCAGAGATGACATCTTGAAGATGACAAAGTATCT	480
QY	934	TTTCTTCGTTAAATGCAATGATGAGACACTCTTGAAAAGCTACTACGCTAGAT	993
DB	481	TTTCTTCGTTAAATGCAATGATGAGACACTCTTGAAAAGCTACTACGCTAGAT	540
QY	994	AGGTGATATGCAAGATGACTTTTGAAGGAGCAAGAGTCTGGTAATTAATGACAC	1053
DB	541	AGGTGATATGCAAGATGACTTTTGAAGGAGCAAGAGTCTGGTAATTAATGACAC	600
QY	1054	ATTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGTCCAAATGGCATTCATCTAT	1113
DB	601	ATTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGTCCAAATGGCATTCATCTAT	660
QY	1114	TGATGACTGCAAGATGATTTCTTAATGACCTTAAAGATCCCTAAATCCGTTGTATA	1173
DB	661	TGATGACTGCAAGATGATTTCTTAATGACCTTAAAGATCCCTAAATCCGTTGTATA	720
QY	1174	TGG 1176	
DB	721	TGG 723	

RESULT 2
BM711125
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BM711125 679 bp mRNA linear EST 28-FEB-2002
UI-E-DX1-agv-1-12-0-UI-E1 UI-E-DX1 Homo sapiens cDNA clone
UI-E-DX1-agv-1-12-0-UI 5', mRNA sequence.
BM711125
BM711125.1 GI:19024383
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 679)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

MEDLINE
8889548
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

source

1. 679
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX1-agv-1-12-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DX1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DX1 is a normalized cDNA library containing the following tissue(s): fetal eyes. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCAGA. This library was created for the program, Gene discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 242 a 133 c 161 g 142 t 1 others

Query Match 19.4%; Score 526; DB 12; Length 679;
Best Local Similarity 99.6%; Pred. No. 4.6e-272;
Matches 676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1964 GAGAAATAGTACATTTTGGAAATGCACTCCGAGTCCGAGAGCGCTGCTTATGG 2023
1 GAGATAGTACATTTTGGAAATGCACTCCGAGTCCGAGAGCGCTGCTTATGG 60
2024 CAATTCACAGGCGGAATGAGAGCGGAAGAGAGATGAGTGTATGATCATATCATC 2083
61 CAATTCACAGGCGGAATGAGAGCGGAAGAGAGATGAGTGTATGATCATATCATC 120
2084 AGACAGATCAAGGCGCTTGTCTAGCTAGTCTTACCAAGAGATTCAGGCAATTAACCTC 2143
121 AGACAGATCAAGGCGCTTGTCTAGCTAGTCTTACCAAGAGATTCAGGCAATTAACCTC 180
2144 TGCATGCGGTGGAACATGCTTCTTAAAGTAACTCTGGAAGTCAAT 2203
181 TGCATGCGGTGGAACATGCTTCTTAAAGTAACTCTGGAAGTCAAT 240
2204 GACACAGACATTTGGAGAACTTCTTAAAGATGATGAGATGCTCTTAAGC 2263
241 GACACAGACATTTGGAGAACTTCTTAAAGATGATGAGATGCTCTTAAGC 300
2264 AAAGAAATGCTCAATGATGACACTTACGCAAGAGTGTGTACGAGACTTCATGCG 2323
301 AAAGAAATGCTCAATGATGACACTTACGCAAGAGTGTGTACGAGACTTCATGCG 360

Qy 2324 CTCATCAACCAACCCCAATCTCAACAGATGATGATGCTTGTGAACAAGTTGGAAAAAG 2383
Db 361 CTCATCAACCAACCCCAATCTCAACAGATGATGATGCTTGTGAACAAGTTGGAAAAAG 420
Qy 2384 GACCGAAGAAACAGTGGCGAAAGGCGAGAGCTATCCCGAGGAAAGTAAAGAAAG 2443
Db 421 GACCGAAGAAACAGTGGCGAAAGGCGAGAGCTATCCCGAGGAAAGTAAAGAAAG 480
Qy 2444 CACTTACAGAAATATGAAAGGTAGAAACAGAGAGACCCAGAAATTTGAGAGGACCC 2503
Db 481 CACTTACAGAAATATGAAAGGTAGAAACAGAGAGACCCAGAAATTTGAGAGGACCC 540
Qy 2504 AGAGTGTCTGAGCTGCTATTAACCTCTAGAAACCTCAACAGATGAGAACTTGCTAGACA 2563
Db 541 AGAGTGTCTGAGCTGCTATTAACCTCTAGAAACCTCAACAGATGAGAACTTGCTAGACA 600
Qy 2564 ATTAAGTGAAGAAACAAATGCAATATACATGAACTTTTTCATGATATGATGATGTT 2623
Db 601 ATTAAGTGAAGAAACAAATGCAATATACATGAACTTTTTCATGATATGATGATGTT 660
Qy 2624 ACAATGCTGGGAAATTCAG 2642
Db 661 ACAATGCTGGGAAATTCAG 679

RESULT 3
BM723807
LOCUS
DEFINITION
UI-E-B01-8ix-o-04-0-UI-r1 UI-E-B01 Homo sapiens cDNA clone
UI-E-B01-8ix-o-04-0-UI 5', mRNA sequence.
ACCESSION
BM723807
VERSION
BM723807.1 GI:19045138
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 702)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

source

1. 702
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-B01-8ix-o-04-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-B01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-B01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares,

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The sequence tag for this library is CCGGATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 248 a 145 c 164 g 145 t
ORIGIN

Query Match 15.4%; Score 418; DB 12; Length 702;
Best Local Similarity 99.6%; Pred. No. 1.1e-213;

Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1972 TAGCAGATTTTGAATGACGTCGAGAGAGCGCTGCTATTGGCAATTCC 2031
DB 52 TAGCAGATTTTGAATGACGTCGAGAGAGCGCTGCTATTGGCAATTCC 111
QY 2032 GAGCGCAATGAAGAGCGAAAGAGATCAGATGATCATATCATCAGACAG 2091
DB 112 GAGCGCAATGAAGAGCGAAAGAGATCAGATGATCATATCATCAGACAG 171
QY 2092 TCAGAGCCTTCTGCTAGTCTACACAGAGATTGAGCAATTACTCTGCATGC 2151
DB 172 TCAGAGCCTTCTGCTAGTCTACACAGAGATTGAGCAATTACTCTGCATGC 231
QY 2152 GGTGGAACATGGTTCATACAACTCTTTAAGTAACTTGAAAGTCAATGACACAG 2211
DB 232 GGTGGAACATGGTTCATACAACTCTTTAAGTAACTTGAAAGTCAATGACACAG 291
QY 2212 GCATTTGGAAGAACTTTCATATAAGTATGATGAGATGCTTAAAGCCAAAGAAAT 2271
DB 292 GCATTTGGAAGAACTTTCATATAAGTATGATGAGATGCTTAAAGCCAAAGAAAT 351
QY 2272 GTCCATATGATGACACCTAGCCAGAGAGTCTGTACAGAGCTTCATGAGCTCATCA 2331
DB 352 GTCCATATGATGACACCTAGCCAGAGAGTCTGTACAGAGCTTCATGAGCTCATCA 411
QY 2332 CCAACCCCAATCTGACACGATGATGATGTTCTGTGAACAAAGTTGGAAGAGGACGGA 2391
DB 412 CCAACCCCAATCTGACACGATGATGATGTTCTGTGAACAAAGTTGGAAGAGGACGGA 471
QY 2392 ACAACCTCGGCAAGGCCAGACATACCCCAAGGAAACATTAAGTAAAGTAACTTAA 2451
DB 472 ACAACCTCGGCAAGGCCAGACATACCCCAAGGAAACATTAAGTAAAGTAACTTAA 531
QY 2452 AGAAATTAAGAAAGTAGAAACAGAGAGACCCAGCAATTT 2491
DB 532 AGAAATTAAGAAAGTAGAAACAGAGAGACCCAGCAATTT 571

RESULT 4
LOCUS CD216784 510 bp mRNA linear EST 20-MAY-2003
DEFINITION EST Y164 Human fibroblasts senescence downregulated subtraced
library Homo sapiens cDNA similar to SEMA3A, mRNA sequence.
ACCESSION CD216784
VERSION CD216784.1 GI:30956759
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 510)
Zhao, L., Zhang, Z., and Tong, T.
Expressed sequence tags from a human subtraced library,
downregulated in senescent fibroblasts
JOURNAL Unpublished
COMMENT Contact: Zhao L

Department of Biochem and Molecular Biology
Peking University, Health Science Center
No.38, Xueyuan Road, Beijing 100083, P R China
Tel: 86 10 62091454
Email: to_zl@pku.edu.cn

FEATURES
source
location/Qualifiers
1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/cell_type="fibroblast"
/cell_line="2BS"
/dev_stage="embryo"
/clone_lib="Human fibroblasts senescence downregulated
subtraced library"
/note="Organ: lung"

BASE COUNT 145 a 111 c 117 g 137 t
ORIGIN

Query Match 15.1%; Score 408; DB 14; Length 510;
Best Local Similarity 99.6%; Pred. No. 2.5e-208;

Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1428 ACAATCCAGTGTTCCTATGAACAAATCCGCCAATATGATCAAAACGATTAATATC 1487
DB 1 ACAATCCAGTGTTCCTATGAACAAATCCGCCAATATGATCAAAACGATTAATATC 60
QY 1488 AATTATACCAATTTGTCGTAACCGAGTGAAGAGATGAGATCAATGATGATGAT 1547
DB 61 AATTATACCAATTTGTCGTAACCGAGTGAAGAGATGAGATCAATGATGATGATGAT 120
QY 1548 TTATCGGAACAGATGTGGAGCCGTTCTTAAAGTATCAATTCCTAAGAGACTGGT 1607
DB 121 TTATCGGAACAGATGTGGAGCCGTTCTTAAAGTATCAATTCCTAAGAGACTGGT 180
QY 1608 ATGATTTAGAGAGGTTCTGCTGAAGAAATGACAGTTTTCGGAACCGACTGCTATT 1667
DB 181 ATGATTTAGAGAGGTTCTGCTGAAGAAATGACAGTTTTCGGAACCGACTGCTATT 240
QY 1668 CAGCATGAGAGTTTCCACTAGCAGCAACATTAATTTGTTCAACCGCTGGGGTTG 1727
DB 241 CAGCATGAGAGTTTCCACTAGCAGCAACATTAATTTGTTCAACCGCTGGGGTTG 300
QY 1728 CCCAGTCCCTTAAACCGGTGATTTAAGGAAAGCGTGTGAGTGTGCTCTG 1787
DB 301 CCCAGTCCCTTAAACCGGTGATTTAAGGAAAGCGTGTGAGTGTGCTCTG 360
QY 1788 CCCGAGCCCTTACTGTGTGAGTGGATGGTTCGATGTTCTGCTAATTTTCCACTGCA 1847
DB 361 CCCGAGCCCTTACTGTGTGAGTGGATGGTTCGATGTTCTGCTAATTTTCCACTGCA 420
QY 1848 AAGAGCGCAAGAGAGCAAGATTAAGAAATGAGAGCCCACTGACTACGTTGAGCT 1907
DB 421 AAGAGCGCAAGAGAGCAAGATTAAGAAATGAGAGCCCACTGACTACGTTGAGCT 480
QY 1908 TACACCATGATATACCATGCGCCACAGCC 1937
DB 481 TACACCATGATATACCATGCGCCACAGCC 510

RESULT 5
LOCUS BF667677 889 bp mRNA linear EST 21-DEC-2000
DEFINITION 60212173P1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810 5',
mRNA sequence.
ACCESSION BF667677
VERSION BF667677.1 GI:11941572
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 889)
Unpublished
COMMENT Contact: Zhao L


```

|||||
Db      225 ATTTGAGAGGACCCGAGAGTGTGAGCTGATTAACCTTGAACCTCAACAAGTA 284
QY      2548 GAACCTTGCTTACATTAATCTGAAAACAAATGCAATATACATGAACCTTTTTCATGG 2607
Db      285 GAACCTTGCTTACATTAATCTGAAAACAAATGCAATATACATGAACCTTTTTCATGG 344
QY      2608 CATTATGTGATGATTTTACATGATGGAATTCAGTGAAGTCCACCAATTAATAATTA 2667
Db      345 CATTATGTGATGATTTTACATGATGGAATTCAGTGAAGTCCACCAATTAATAATTA 404
QY      2668 ATCCATGATGATTAACCTTTCTTAATAGGCT 2694
Db      405 ATCCATGATGATTAACCTTTCTTAATAGGCT 431

RESULT 7
LOCUS    BX283116                      382 bp      mRNA      linear      EST 04-MAR-2003
DEFINITION
IMAGE:4278810, mRNA sequence.
ACCESSION
BX283116
VERSION  EX283116.1  GI:28847570
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 382)
Ebert,L., Heil,O., Hennig,S., Neubert,P., Patsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Human Unigeneset - RZPD3
Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 560, D-69120 Heidelberg, Germany
RZPD; IMAGE958A191103.
RZPDLIB; I.M.A.G.E. CDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi?responseLibNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMV-M13u, Primer sequence: CGTTGTAAACGACGGCCAGT.
Location/Qualifiers
1..382
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="IMAGE958A191103 ; IMAGE:4278810"
/risue_type="primitive neuroectoderm"
/lab host="DH10B (TI phage-resistant)"
/clone.lib="NIH MGC 56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
Stit1 (ggccgcctcgcc); Site_2: Stit1 (ggccattggc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCGCAGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 clones
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 106 a 85 c 94 g 97 t
ORIGIN
Query Match 14.1%; Score 382; DB 13; Length 382;

```

```

Best Local Similarity 100.0%; Pred. No. 2.6e-194;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1595 AAGGAGACTTGTATGATTTTAAAGAGCTTCTGCTGGAAGAAATGACACTTTTGGGAA 1654
Db      1 AAGGAGACTTGTATGATTTTAAAGAGCTTCTGCTGGAAGAAATGACACTTTTGGGAA 60
QY      1655 CCGACTGCTATTTCAGCAATGAGAGCTTTCACATAAGCAGCAACATATATATTGGTTCA 1714
Db      61 CCGACTGCTATTTCAGCAATGAGAGCTTTCACATAAGCAGCAACATATATATTGGTTCA 120
QY      1715 ACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGTATATTAGGGAAAGCGTGTCT 1774
Db      121 ACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGTATATTAGGGAAAGCGTGTCT 180
QY      1775 GAGGTGTCGTCGCCCGAGACCCCTTACATGCTGTGGAGTGTTCGCAATGTTCTGCTAT 1834
Db      181 GAGGTGTCGTCGCCCGAGACCCCTTACATGCTGTGGAGTGTTCGCAATGTTCTGCTAT 240
QY      1835 TTTCCCACTGCAAGAGAGAGCAAGAGAGCAAGATATTAAGAAATGAGAGCCACTGACT 1894
Db      241 TTTCCCACTGCAAGAGAGAGCAAGAGAGCAAGATATTAAGAAATGAGAGCCACTGACT 300
QY      1895 CACTGTTGACTTACACCATATATATCCATGCGCACAGCCTTGAAGAGAGATCATC 1954
Db      301 CACTGTTGACTTACACCATATATATCCATGCGCACAGCCTTGAAGAGAGATCATC 360
QY      1955 TATGGTGTAGAAATATGATGCA 1976
Db      361 TATGGTGTAGAAATATGATGCA 382

RESULT 8
HSM084323/C
ID      HSM084323  standard; RNA; EST; 721 BP.
AC      BX510254;
SV      BX510254.1
DT      09-MAY-2003 (Rel. 75, Created)
DT      09-MAY-2003 (Rel. 75, Last updated, Version 1)
DE      Homo sapiens mRNA; EST DKFZp686L199_s1 (from clone DKFZp686L1998)
XX      EST; expressed sequence tag.
XX      Homo sapiens (human)
XX      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX      Eutheria; Primates; Catarrhini; Homnidae; Homo.
XX      [1]
XX      1-721
XX      Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
XX      Han M., Wiemann S.;
XX      Submitted (09-MAY-2003) to the EMBL/Genbank/DBJ databases.
XX      MIPS, Ingolstaedter Landstr.1, D-85764 Neubherberg, GERMANY
XX      This is the 3' sequence of the clone insert
XX      CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
XX      CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
XX      CC sequenced by Olegen (Hilden/Germany) within the cDNA sequencing
XX      CC consortium of the German Genome Project.
XX      CC r1 sequence also available.
XX      CC This clone (DKFZp686L1998) is available at the RZPD in Berlin.
XX      CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
XX      CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
FH      Key      Location/Qualifiers
FH      source      1..721
FH      FT          /db_xref="taxon:9606"

```



```

FT      /mol_type="mRNA"
FT      /organism="Homo sapiens"
FT      /clone="DKFZp686L198"
FT      /clone_1b="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT      DH10B; sites SfiI + SfiI"
FT      /dev_stage="adult"
FT      /tissue_type="CDNA-collection"
XX      SQ      Sequence 721 BP; 192 A; 141 C; 121 G; 267 T; 0 other;

Query Match      12.0%; Score 324; DB 2; Length 721;
Best Local Similarity 100.0%; Pred. No. 6.8e-163;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2377 GAAAGGAGCCGAAACCAAGTCGCGAAGGCGAGACATACCCGAGAACATACAA 2436
DB      690 GAAAGGAGCCGAAACCAAGTCGCGAAGGCGAGACATACCCGAGAACATACAA 631
QY      2437 ATGGAAGCCTTACAGAAATTAAGAAAGTGAAGACAGAGAGCCAGAAATTTGAGG 2496
DB      630 ATGGAAGCCTTACAGAAATTAAGAAAGTGAAGACAGAGAGCCAGAAATTTGAGG 571
QY      2497 GGCACCCGAGAGTGTCTGAGCTGAGCTTACCTCTAGAAACCTCAACAAAGTAACTGC 2556
DB      570 GGCACCCGAGAGTGTCTGAGCTGAGCTTACCTCTAGAAACCTCAACAAAGTAACTGC 511
QY      2557 CTAGCAATTAATCTGAGAAACCAATGCAATATATATATATATATATATATATAT 2616
DB      510 CTAGCAATTAATCTGAGAAACCAATGCAATATATATATATATATATATATATATAT 451
QY      2617 GATGTTTACATATGTTGGGAAATTCAGCTGAGTTCACCAATTAATTAATTCATG 2676
DB      450 GATGTTTACATATGTTGGGAAATTCAGCTGAGTTCACCAATTAATTAATTCATG 391
QY      2677 TAACCTTCTAATAGCTTTT 2700
DB      390 TAACCTTCTAATAGCTTTT 367

RESULT 9
LOCUS      BM670627/c      670 bp      mRNA      linear      EST 27-FEB-2002
DEFINITION      UI-E-DX1-egv-1-12-0-UI.s1 UI-E-DX1 Homo sapiens cDNA clone.
ACCESSION      BM670627
VERSION      BM670627.1 GI:18980524
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      1 (bases 1 to 670)
AUTHORS      Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
PUBMED      8889548
COMMENT      Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-25, >AT_rich#low_complexity (matched complement)

```

```

FEATURES
source
Seq primer: M13 Forward
PolyA=yes.
Location/Qualifiers
1..670
/mol_type="mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-DX1-egv-1-12-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1b="UI-E-DX1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-DX1 is a normalized cDNA library containing the
following tissue(s): fetal eyes. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGAAATCAAGA. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_LIB=UI-E-DX1
TAG_TISSUE=human fetal eyes
TAG_SEQ=AGAAATCAAGA"
BASE COUNT      181 a      128 c      115 g      246 t
ORIGIN

Query Match      11.6%; Score 315; DB 12; Length 670;
Best Local Similarity 100.0%; Pred. No. 4.9e-158;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2395 ACGTCGGCAAGGCGAGACATACCCGAGGAGAGTAAATGAGACATTAACAAGA 2454
DB      670 ACGTCGGCAAGGCGAGACATACCCGAGGAGAGTAAATGAGACATTAACAAGA 611
QY      2455 AAATAGAAAGGTAGAAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTCTG 2514
DB      610 AAATAGAAAGGTAGAAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTCTG 551
QY      2515 AGCTGCATTACCTTAGAAACCTTAACAGTAGAAACCTGCTAGACATTAACGTGAAA 2574
DB      550 AGCTGCATTACCTTAGAAACCTTAACAGTAGAAACCTGCTAGACATTAACGTGAAA 491
QY      2575 AACAAATGCAATTAATCAATGAACTTTTCAAGGATTAATGATGATTAATGATGAGG 2634
DB      490 AACAAATGCAATTAATCAATGAACTTTTCAAGGATTAATGATGATTAATGATGAGG 431
QY      2635 AAATTCAGCTAGTTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2694
DB      430 AAATTCAGCTAGTTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 371
QY      2695 TTTTTCCTAATACC 2709
DB      370 TTTTTCCTAATACC 356

RESULT 10
LOCUS      BF667143      839 bp      mRNA      linear      EST 21-DEC-2000
DEFINITION      602121928F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278967 5',
ACCESSION      BF667143
VERSION      BF667143.1 GI:11941038
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```



```

Db      450 TGGATTTTACATGGTGGAAATTCAGCTGAGTCCACCAATTATTAATTCATG 391
Qy      2675 AGTACTTTCCTAATAGGCTTTTTCCTAATACC 2709
Db      390 AGTAACTTTCCTAATAGGCTTTTTCCTAATACC 356

RESULT 12
BF700780      884 bp      mRNA      linear      EST 22-DEC-2000
LOCUS      602128585F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285532 5',
DEFINITION      mRNA sequence.
ACCESSION      BF700780
VERSION      BF700780.1 GI:11986188
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgabds-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: CLONETECH Laboratories, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM1120 row: 1 column: 21
              High quality sequence stop: 676.
              Location/Qualifiers
                1..884
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:4285532"
                  /tissue_type="primitive neuroectoderm"
                  /lab_host="DH10B (T1 phage-resistant)"
                  /clone_1lb="NIH_MGC_56"
                  /note="Organ: brain; Vector: pDNR-LIB (Clontech); site_1:
                  SfiI (ggccgctcgcc); Site 2: SfiI (ggccatcgcc);
                  Double-stranded cDNA was prepared from cell line RNA. 5'
                  and 3' adaptors were used in cloning as follows: 5'
                  adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor
                  sequence: 5'-ATTCTAGAGCGCGAGCGCGACATG-dT(30)BN-3'
                  (where B = A, C, or G and N = A, C, G, or T). Average
                  insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
                  contained inserts by PCR. This library was enriched for
                  full-length clones and was constructed by Clontech
                  Laboratories (Palo Alto, CA)."
BASE COUNT      280 a 185 c 225 g 194 t
ORIGIN
Query Match      10.1%; Score 274; DB 10; Length 884;
Best Local Similarity 100.0%; Pred. No. 7.6e-136; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 0;
Qy      1916 GATAATCACCATGCGCACAGCCCTGAAGAGAGAAATCATCTATGTTAGAGATAGTAC 1975
Db      52 GATATACACCATGCGCACAGCCCTGAAGAGAGAAATCATCTATGTTAGAGATAGTAC 111
Qy      1976 ACATTTTGAATGACGTCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAG 2035
Db      112 ACATTTTGAATGACGTCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAG 171
Qy      2036 CGAAATGAGAGGAAAGAGAGATCAGAGTGAATCATATCTATCTGAGACAGATCA 2095
Db      172 CGAAATGAGAGGAAAGAGAGATCAGAGTGAATCATATCTATCTGAGACAGATCA 231

```

```

Qy      2096 GGCCTTTCGTCAGTACTCTACACAGAGATTCAGGCAATTCCTGCGATCGGTG 2155
Db      232 GGCCTTTCGTCAGTACTCTACACAGAGATTCAGGCAATTCCTGCGATCGGTG 291
Qy      2156 GAACATGGGTCATCAAACTCTCTTAAGGTA 2189
Db      292 GAACATGGGTCATCAAACTCTCTTAAGGTA 325

RESULT 13
HSM084322      standard, RNA; EST; 535 BP.
ID      HSM084322
AC      BX491494;
XX      BX491494.1
SV      BX491494.1
DT      09-MAY-2003 (rel. 75, Created)
DT      09-MAY-2003 (rel. 75, Last updated, Version 1)
DE      Homo sapiens mRNA; EST DKFZp686L1998_r1 (from clone DKFZp686L1998)
XX      EST; expressed sequence tag.
KW      EST; expressed sequence tag.
OS      Homo sapiens (human)
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC      Eutheria; Primates; Catarrhini; Homiidae; Homo.
XX      [1]
RN      1-535
RP      Bahr A., Lauber J., Mewes H.W., Well B., Amid C., Osanger A., Fobo G.,
RA      Han M., Wiemann S.;
RT      Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.
RL      MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY
XX      This is the 5' sequence of the clone insert
CC      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC      Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC      sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
CC      consortium of the German Genome Project.
CC      SI sequence also available.
CC      This clone (DKFZp686L1998) is available at the RZPD in Berlin.
CC      Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC      14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX      Key      Location/Qualifiers
FH      1..535
FH      source
FH      /db_xref="taxon:9606"
FH      /mol_type="mRNA"
FH      /organism="Homo sapiens"
FH      /clone_1lb="DKFZp686L1998"
FH      /note="synonym: hlc33). Vector pSport1_Sfi; host
FH      DH10B; sites SfiI + SfiIIB"
FH      /dev stage="adult"
FH      /tissue_type="cDNA-collection"
SQ      Sequence 535 BP; 135 A; 108 C; 119 G; 173 T; 0 other;
Query Match      9.8%; Score 265; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 4.8e-111; Indels 0; Gaps 0;
Matches 265; Conservative 0; Mismatches 0;
Qy      118 TTTCAGATGTTGAACTTCTCTGCGCGACACATACAGAGAAAGACTAAAGACGAA 177
Db      271 TTTCAGATGTTGAACTTCTCTGCGCGACACATACAGAGAAAGACTAAAGACGAA 330
Qy      178 AGGAGCTTACAGCGCTGTCAGAGATGGGCTGTGTAATCTGATGTTGCTCTTTCTGAGG 237
Db      331 AGGAGCTTACAGCGCTGTCAGAGATGGGCTGTGTAATCTGATGTTGCTCTTTCTGAGG 390
Qy      238 AGTATTAAGAGAGAGAAAGATCAATCAAGATGGAAGAAACAATGTGCCAAGGCTGA 297

```


ORIGIN

Query Match 8.0%; Score 218; DB 10; Length 355;
 Best Local Similarity 100.0%; Pred. No. 1.1e-105; Indels 0; Gaps 0;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 TGAGGACATATTTTAACTGAGAACTCAGATTTTGAAAACGGCCGTGGAGAGTCC 708
 Db 355 TGAGGACATATTTTAACTGAGAACTCAGATTTTGAAAACGGCCGTGGAGAGTCC 296
 QY 709 AATATGACCTTAAGCTCTGACAGCATCCCTTTAATAGTGAATTTACTCTGGAAC 768
 Db 295 AATATGACCTTAAGCTCTGACAGCATCCCTTTAATAGTGAATTTACTCTGGAAC 236
 QY 769 TCGAGTGAATTTATGAGGAGAGCTTGTCTTCCGAACTCTTGGGACACACACC 828
 Db 235 TCGAGTGAATTTATGAGGAGAGCTTGTCTTCCGAACTCTTGGGACACACACC 176
 QY 829 AATCAGACAGACAGCATGATTTCCAGTGGCTCAATG 866
 Db 175 AATCAGACAGACAGCATGATTTCCAGTGGCTCAATG 138

RESULT 16 231 bp mRNA linear EST 06-SEP-1995
 LOCUS T29595 EST6074 Human lung Homo sapiens CDNA 5' end similar to semaphorin
 DEFINITION III (HT:3407), mRNA sequence.

ACCESSION T29595
 VERSION T29595.1 GI:611693
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 1 (bases 1 to 231)

TITLE Initial Assessment of Human Gene Diversity and Expression Patterns
 JOURNAL Nature 377, 3-174 (1995)
 MEDLINE 96026280
 PUBMED 7566098

COMMENT Contact: Venter, JC
 The Institute for Genomic Research
 932 Clopper Rd, Gaithersburg, MD 20878
 Tel: 3018699056
 Fax: 3018699423

FEATURES
 source
 1..231
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):106785"

/db_xref="taxon:9606"
 /clone_11b="Human Lung"
 /note="Organ: Lung"
 BASE COUNT 65 a 48 c 56 g 62 t
 ORIGIN

Query Match 7.9%; Score 213; DB 14; Length 231;
 Best Local Similarity 100.0%; Pred. No. 4.9e-103;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1123 GCAGATGATTTCTTAATGAACCTTAAGATCTTAAATCCAGTGTATGAGTGT 1182
 Db 1 GCAGATGATTTCTTAATGAACCTTAAGATCTTAAATCCAGTGTATGAGTGT 60
 QY 1183 TACGACTTCAGTAACTTTTCAAGGAGTACCGGTGTATGATATGAGTGT 1242
 Db 61 TACGACTTCAGTAACTTTTCAAGGAGTACCGGTGTATGATATGAGTGT 120
 QY 1243 GAGAAAGGTGTTCTTGTCCATATGCCCACAGGATGACCACTATCATATGAGTCC 1302
 Db 121 GAGAAAGGTGTTCTTGTCCATATGCCCACAGGATGACCACTATCATATGAGTCC 180
 QY 1303 TTATCAAGAAAGAGTCCCTTATCCAGGCTCAGG 1335
 Db 181 TTATCAAGAAAGAGTCCCTTATCCAGGCTCAGG 213

RESULT 17 328 bp mRNA linear EST 17-DEC-1995
 LOCUS A1671991/c wb72c04.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:2311206 3'
 DEFINITION similar to TR:Q14563 Q14563 SEMAPHORIN-III., mRNA sequence.

ACCESSION A1671991
 VERSION A1671991.1 GI:4851722
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 1 (bases 1 to 328)
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNI, at:
 www.bio.lnlnl.gov/bbrp/image/image.html
 Insert length: 358 Std Error: 0.00
 Seq primer: -40bp from Gibco.

FEATURES
 source
 1..328
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_11b="NCI CGAP GC6"
 /note="Vector: pTR13D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
 from the normalized library NCI-CCAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtracting hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clonoids

BASE COUNT 1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

Query Match 7.5%; Score 203; DB 9; Length 328;
Best Local Similarity 100.0%; Pred. No. 1,4e-97;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 807 GAACCTCTGGGACACCCACCAATGAGACAGACAGACAGATTCAGAGGCTGCAATG 866
DB 328 GAACCTCTGGGACACCCACCAATGAGACAGACAGACAGATTCAGAGGCTGCAATG 269
Y 867 ATCCAAAGTTCATTAAGTCCACCTCATCTCAGAGAGTGAACAATCTGAAGATGACAAG 926
DB 268 ATCCAAAGTTCATTAAGTCCACCTCATCTCAGAGAGTGAACAATCTGAAGATGACAAG 209
Y 927 TATACCTTTTCTCCGAGAAAATGCAATAGAGAACTCTGAGAAAAGCTACTACG 986
DB 208 TATACCTTTTCTCCGAGAAAATGCAATAGAGAACTCTGAGAAAAGCTACTACG 149
Y 987 CTGAATAGGTCAGATATGCAAG 1009
DB 148 CTGAATAGGTCAGATATGCAAG 126

RESULT 18 210 bp mRNA linear EST 17-APR-1995
R19784
LOCUS R19784
DEFINITION Y928612.r1 Soares infant brain INIB Homo sapiens cDNA clone
sequence.
IMAGE:33664.5' similar to SP:A49069 A49069 COLLA5IN - ; mRNA

ACCESSION R19784
VERSION R19784.1 GI:774418
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 210)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hultman, M., Kucada, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, B., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
R., Williamson, A., Wohlmann, P. and Wilson, R.
The Washu-Merck EST Project
Unpublished

TITLE Contact: Wilson RK
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

Insert Size: 1370
High quality sequence stops: 127 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1370 Std Error: 0.00
Seg primer: M13Rpl
High quality sequence stop: 127.
Location/Qualifiers

FEATURES
source 1..210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:406011"
/db_xref="taxon:9606"
/clone="IMAGE:33664"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain INIB"
/note="Organ: whole brain; Vector: Laddid BA; Site: 1: Not
1; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo (dT) primer [5'

BASE COUNT 65 a 44 c 53 g 46 t
Query Match 7.2%; Score 196; DB 14; Length 210;
Best Local Similarity 100.0%; Pred. No. 7.3e-94;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2000 TCGCAGAGCCCTGCTTATTCAGATTCAGAGCCGAAATGAAGGAAAGAG 2059
DB 1 TCGCAGAGCCCTGCTTATTCAGATTCAGAGCCGAAATGAAGGAAAGAG 60
Y 2060 ATCAGAGTGCATATTCATTCAGAGACATCAAGGCTTCTGCTAGTACTACAA 2119
DB 61 ATCAGAGTGCATATTCATTCAGAGACATCAAGGCTTCTGCTAGTACTACAA 120
Y 2120 CAGAGGATTCAGGCAATTAACCTGCGCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
DB 121 CAGAGGATTCAGGCAATTAACCTGCGCATGCGGTGGAACATGGGTTCATACAACTCTT 180
Y 2180 CTTAAGGTAAACCTG 2195
DB 181 CTTAAGGTAAACCTG 196

RESULT 19 914 bp mRNA linear EST 23-MAY-2003
AL529332
LOCUS AL529332
DEFINITION AL529332 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CSODD002YNI6 5-PRIME, mRNA sequence.
IMAGE:33664.5' similar to SP:A49069 A49069 COLLA5IN - ; mRNA

ACCESSION AL529332
VERSION AL529332.2 GI:31067175
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 914)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On Feb 13, 2001 this sequence version replaced gi:12792825.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4349.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODD002D0808P1&cluster=4349.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CSODD002D0808P1.

Insert Size: 1370
High quality sequence stops: 127 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1370 Std Error: 0.00
Seg primer: M13Rpl
High quality sequence stop: 127.
Location/Qualifiers

FEATURES
source 1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODD002YNI6"
/cissue="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and BclI V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 279 a 198 c 183 g 253 t
ORIGIN

Query Match 7.2%; Score 194; DB 9; Length 914;
 Best Local Similarity 100.0%; Pred. No. 1.3e-92;
 Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCCGACAAATACAGAGAGAAAGACTAAAGACGAA 177
 DB 479 TTTCAGATTGTTGAACCTCTCTGCGCCGACAAATACAGAGAGAAAGACTAAAGACGAA 538
 QY 178 AGGACCTACAGAGCTGAGAGATGGGCTGTTACAGAGATGCTGCTTTCTGAGGAG 237
 DB 539 AGGACCTACAGAGCTGAGAGATGGGCTGTTACAGAGATGCTGCTTTCTGAGGAG 598
 QY 238 AGTATTACTTACAGAGAGCAAACTATACAGATGGGAGAAACATATGCTCCAGGCTGAA 297
 DB 599 AGTATTACTTACAGAGAGCAAACTATACAGATGGGAGAAACATATGCTCCAGGCTGAA 658
 QY 298 ATTATCTCTCAAG 311
 DB 659 ATTATCTCTCAAG 672

RESULT 20
 BE552070/c 325 bp mRNA linear EST 10-AUG-2000
 LOCUS hy03d11.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:3196245 3'
 DEFINITION similar to TR:Q14563 Q14563 SEMAPHORIN-III.; mRNA sequence.
 ACCESSION BE552070
 VERSION BE552070.1 GI:9793762
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 TITLE 1 (bases 1 to 325)
 JOURNAL NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 COMMENT Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seg primer: -40up from Gibco.

FEATURES
 source Location/Qualifiers
 1..325

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3196245"
 /issue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_lib="NCI-CCAP GC6"
 /note="Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CCAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clone IDs 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 81 a 63 c 67 g 114 t
 Query Match 5.0%; Score 135; DB 10; Length 325;

Best Local Similarity 100.0%; Pred. No. 7.6e-61;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 807 GAACCTTGGGACACACCAACCAACAGACAGACACATTCAGGCTCATG 866
 DB 325 GAACCTTGGGACACACCAACCAACAGACAGACACATTCAGGCTCATG 266
 QY 867 ATCCAAAGTTCATTAGTCCCACTCATCTCAGAGATGACCAATCTTAAGTGAAGA 926
 DB 265 ATCCAAAGTTCATTAGTCCCACTCATCTCAGAGATGACCAATCTTAAGTGAAGA 206
 QY 927 TATACCTTTTCTTCC 941
 DB 205 TATACCTTTTCTTCC 191

RESULT 21
 AA424316 430 bp mRNA linear EST 16-OCT-1997
 LOCUS zv90d08.r1 Soares NHMPu_S1 Homo sapiens CDNA clone IMAGE:767055
 DEFINITION 5', mRNA sequence.
 ACCESSION AA424316
 VERSION AA424316.1 GI:2103286
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 TITLE 1 (bases 1 to 430)
 JOURNAL Hillier, L., Allen, M., Bowles, J., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Lennon, G., Merrin, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 Washu-Merck EST Project 1997
 COMMENT Unpublished
 CONTACT: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seg primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 413.
 FEATURES
 source Location/Qualifiers
 1..430

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:767055"
 /issue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares NHMPu_S1"
 /note="Organ: mixed (see below); Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHn, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 152 a 82 c 78 g 118 t
 Query Match 4.4%; Score 120; DB 9; Length 430;
 Best Local Similarity 100.0%; Pred. No. 1e-52;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	Db	QY	Db
2485	CGAATTTCAGAGGGGACCCAGSAGGTCTGAGCTGCATTACTCTAGAAA	2544	CGTCAAAACA
37	CGAATTTCAGAGGGGACCCAGSAGGTCTGAGCTGCATTACTCTAGAAA	96	CTCAAAACA
2545	GTAGAAATCTTGCTGACATATATCTGGAAAAACAAATGCATATATCA	2604	CTTTTTC
97	GTAGAAATCTTGCTGACATATATCTGGAAAAACAAATGCATATATCA	156	GAATCTTTTTC

RESULT	22
BUI66364	
LOCUS	BUI66364
DEFINITION	BUI66364 896 bp mRNA linear EST 04-SEP-2002 AGNCOURT 7953921 NIH_MGC_68 Homo sapiens CDNA clone IMAGE:6010676 5', mRNA sequence.

RESULT	23
BE033028	
LOCUS	
DEFINITION	133244 MRC 1Ptg Sus scrofa CDNA 5' mRNA linear EST 09-JUL-2000
ACCESSION	BE033028
VERSION	BE033028.1 GI:8327974
KEYWORDS	EST.
SOURCE	Sus.scrofa (pig)

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM12287 row: h column: 06
 High quality sequence stop: 500.

FEATURES

source

1. .1005
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5560757"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 88"
 /note="Organ: small_intestine; Vector: PCMV-SPORT6;
 Site_1: NotI; Site_2: SalI; cloned unidirectionally;
 oligo-dT primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 262 a 283 c 138 g 322 t
 ORIGIN

Query Match 2.5%; Score 68; DB 12; Length 1005;
 Best Local Similarity 100.0%; Pred. No. 1.7e-24;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 AAAAGATTGTCGCGCATCTTACACGAAAGAGTGAATGACAGTGGCTGGAAGA 524
 |||||
 1 AAAAGATTGTCGCGCATCTTACACGAAAGAGTGAATGACAGTGGCTGGAAGA 60

QY 525 ACATCCTG 532
 |||||
 61 ACATCCTG 68

RESULT 25 AG157495 677 bp DNA linear GSS 09-JAN-2002
 LOCUS AG157495
 DEFINITION Pan troglodytes DNA, clone: RP43-022G24.T7, genomic survey
 sequence.
 ACCESSION AG157495
 VERSION AG157495.1 GI:16687173
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of Library RPCT-43
 JOURNAL Unpublished
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission

COMMENT Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenho-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbesc@sc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RPCT-43 This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.
 PRIMERS
 Sequencing: T7

LIBRARY Vector : PBAC63.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI
 Location/Qualifiers

source

1. .677
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-022G24.T7"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RPCT-43 Chimpanzee Male BAC Library"
 BASE COUNT 243 a 106 c 91 g 236 t 1 others
 ORIGIN

Query Match 2.5%; Score 67; DB 29; Length 677;
 Best Local Similarity 100.0%; Pred. No. 5.2e-24;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 AAGCCACAGACGACAGATATTAAGATGAGACCCAGTGCACCTTACAGCTTA 1909
 |||||
 312 AAGCCACAGACGACAGATATTAAGATGAGACCCAGTGCACCTTACAGCTTA 371

QY 1910 CACCATG 1916
 |||||
 372 CACCATG 378

RESULT 26 BI183160 604 bp mRNA linear EST 10-JUL-2001
 LOCUS BI183160
 DEFINITION UNI-P-FN-bl-a-10-0-UNL.s1 UNI-P-FN Sus scrofa cDNA clone
 UNI-P-FN-bl-a-10-0-UNL 3', mRNA sequence.
 ACCESSION BI183160
 VERSION BI183160.1 GI:14657569
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 604)
 Caetano, A.R., Johnson, R.K. and Pomp, D.
 Generation and sequence characterization of a normalized cDNA
 library from swine ovarian follicles
 Mamm. Genome 14 (1), 65-70 (2003)

JOURNAL MEDLINE
 PUBMED
 COMMENT Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu

Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message.
 Seq primer: M13 -29
 POLY-A-NO.

FEATURES

source

1. .604
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /strain="University of Nebraska, Lincoln Swine Selection
 Lines"
 /db_xref="taxon:9823"
 /clone="UNI-P-FN-bl-a-10-0-UNL"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UNI-P-FN"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UNI-P-FN
 library is a normalized library representing porcine
 ovarian follicles, ranging between 2.0 to 10.0 mm in
 diameter, collected during 7 days of the follicular phase
 of the pig estrous cycle. This library was derived from
 the library UNI-P-F2. The tag is a string of 5-6
 nucleotides present between the Not I site and the

BASE COUNT

180 a 132 c 130 g 162 t

ORIGIN

oligo-dt track. The library was constructed as described by Bonaldo, Lemon and Soares, Genome Research 6: 791-806, 1996.

TAG SEQ=None found"

Query Match 2.2%; Score 59; DB 12; Length 604;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 770 GCACGCTGATTTTATGGGCGAGACTTGTCTATCTTCGAACTCTTGGGCGACCC 828
DB 234 GCACGCTGATTTTATGGGCGAGACTTGTCTATCTTCGAACTCTTGGGCGACCC 292

RESULT 27

LOCUS

BX282800 598 bp mRNA linear EST 04-MAR-2003
BX282800 Soares NHPu S1 Homo sapiens cDNA clone IMAGE998N161887 ;
IMAGE:767055, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGE998N161887.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD (RZPDLIB No.972)
http://www.rzpd.de/Cloncards/cgi-bin/showlib.pl/cgi/respenserlibno=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r. Primer sequence: TTTCAACAGGAAACGCTATGAC.
Location/Qualifiers

FEATURES

source

1..598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998N161887 ; IMAGE:767055"
/tissue type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab host="DH10B"

/clone_lib="Soares NHPu S1"
/note="Organ: mixed (see below); Vector: pTT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NDM, pregnant uterus
NHPu, and fetal heart NHPu) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT

191 a 101 c 105 g 164 t 37 others

ORIGIN

Query Match 2.1%; Score 57; DB 13; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2522 TTACCTCTGAACCTCAACAGTAGAATCTGCTACACATACCTGAAACA 2578
DB 75 TTACCTCTGAACCTCAACAGTAGAATCTGCTACACATACCTGAAACA 131

RESULT 28

LOCUS

AO533647 576 bp DNA linear GSS 18-MAY-1999
RPCT-11-35304.TJ RPCT-11 Homo sapiens genomic clone RPCT-11-35304,
genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other GSSs: RPCT-11-35304.TJ
Map Building
Unpublished
Use of BAC End Sequences from Library RPCT-11 for Sequence-Ready
Map Building
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@ig.org
Clones are derived from the human BAC library RPCT-11. For BAC
library availability, please contact Pleher de Jong
(pleher@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genes (http://info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: Sp6
Classes: BAC ends.
Location/Qualifiers

FEATURES

source

1..576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:763507"
/db_xref="taxon:9606"
/clone="RPCT-11-35304"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCT-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCT11 Human Male BAC Library"
BASE COUNT 202 a 86 c 90 g 197 t 1 others
ORIGIN

Query Match 2.1%; Score 56; DB 28; Length 576;
Best Local Similarity 100.0%; Pred. No. 4.4e-18;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 399 GTAGGCTGATTTGGAGCAAGATTCATATTTTCATGACCTGGTTATATC 454
DB 576 GTAGGCTGATTTGGAGCAAGATTCATATTTTCATGACCTGGTTATATC 521

RESULT 29

LOCUS

CA872840 559 bp mRNA linear EST 20-DEC-2002
K0920H08-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
Library (long) Mus musculus cDNA clone NIA:K0920H08 IMAGE:30086203
5', mRNA sequence.

ORIGIN

ACCESSION CA872840
 VERSION CA872840.1 GI:27324389
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 559)
 AUTHORS Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Alpha, K., Vescevi, A.L. and Ko, M.S.H.
 TITLE Systematic Analyses of NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)
 JOURNAL Unpublished
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: K0920 row: H column: 08
 Seq primer: -21M13 Reverse
 High quality sequence stop: 559
 POLYA=No.

FEATURES
 source
 1..559
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="CD1"
 /db_xref="taxon:10090"
 /clone="NIA:K0920H08 IMAGE:30086203"
 /dev_stage="Adult"
 /lab_host="DHI08"
 /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen). Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544191]). Total RNAs were obtained from Dr. Angelo L. Vescevi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGCGCGCCGCTTTTCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-SalI, purified by phenol/chloroform, and separated from free linkers by centrifugation 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer SalI-8. The products were purified by phenol/chloroform and centrifugation 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DHI08 B. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.8 kb. The library was constructed by Yulan Piao."

BASE COUNT
 ORIGIN 166 a 115 c 126 g 152 t

Query Match 1.9%; Score 52; DB 14; Length 559;
 Best Local Similarity 100.0%; Pred. No. 6.3e-16;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 GCTATTTCAGATTGTTGAATCTTCTGCGCGCATATACGAGAGAG 164
 ||||||||||||||||||||||||||||||||||||||||||||||||
 DB 3 GCTATTTCAGATTGTTGAATCTTCTGCGCGCATATACGAGAGAG 54

RESULT 30
 AA716143 394 bp mRNA linear EST 29-DEC-1997
 LOCUS AA716143/c
 DEFINITION Zg60d02.s1 Soares_fetal_heart_NbH19w Homo sapiens cDNA clone

IMAGE:397731 3', mRNA sequence.
 AA716143
 VERSION AA716143.1 GI:2728417
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 394)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellendberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 363.

FEATURES
 source
 1..394
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1305554"
 /db_xref="taxon:9606"
 /clone="IMAGE:397731"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DHI08 (ampicillin resistant)"
 /clone_lib="Soares_fetal_heart_NbH19w"
 /note="Organ: heart; Vector: pRT3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TCTTTCAGATCGAAGTGGAGCGCGCGCATTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaudo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbH19w."

BASE COUNT
 ORIGIN 128 a 69 c 62 g 135 t

Query Match 1.9%; Score 51; DB 9; Length 394;
 Best Local Similarity 100.0%; Pred. No. 2e-15;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2659 ATAAATTAATCAGATGATTAATCTTCTGCGCGCATATACG 2709
 ||||||||||||||||||||||||||||||||||||||||||||||||
 DB 394 ATAAATTAATCAGATGATTAATCTTCTGCGCGCATATACG 344

RESULT 31
 BH265538 773 bp DNA linear GSS 30-NOV-2001
 LOCUS BH265538
 DEFINITION CH230-43J18.TU CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-43J18, genomic survey sequence.
 ACCESSION BH265538
 VERSION BH265538.1 GI:17175582
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 773)
 AUTHORS Zhao, S., Shetty, J., Shateman, S., Tesgaye, G., Geer, K., Shvartsbeyn, A., Gebregiorgis, B., Overton, L., Russell, D., Chen, D., Shyng, F., de Jong, P. and Fraser, C.M.
 TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 JOURNAL Unpublished
 COMMENT Other_GSSS: CH230-43J18.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (http://www.chori.org/bacpac/rat230.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering/information.htm). BAC end
 plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 43 row: J column: 18
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..773
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SsNHd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-43J18"
 /sex="Female"
 /cell_type="Brain"
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SsNHd/MCW) BAC library produced by
 Pieter de Jong"

BASE COUNT 240 a 159 c 157 g 217 t

ORIGIN

Query Match 1.8%; Score 50; DB 28; Length 773;
 Best Local Similarity 100.0%; Pred.No. 8.3e-15;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2291 AGCCAGAGTCTGTACAGAGACTTCATGACAGCTATCAACCCCA 2340
 |||||
 571 AGCCAGAGTCTGTACAGAGACTTCATGACAGCTATCAACCCCA 620

Db

RESULT 32
 BG315503 124 bp mRNA linear EST 28-FEB-2002
 LOCUS BG315503
 DEFINITION POS 0.131 Human THP1 cell line library Homo sapiens cDNA, mRNA
 sequence.
 ACCESSION BG315503
 VERSION BG315503.1 GI:18998375
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 124)
 Andersson, T., Borzang, S., Larsson, M., Thelin, A., Ekstrand-Hammarstrom, B., Wirta, V., Wenborg, A., Lundberg, J. and Odeberg, J.
 Identification of candidate genes in atherosclerosis - Virtual chip
 analysis in RDA based transcript profiling of monocyte/macrophage
 response to oxidised LDL
 Unpublished
 Contact: Andersson Tove
 Department of Biotechnology
 KTH
 Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
 Tel: +46 8 790 71 29
 Fax: +46 8 245452

FEATURES
 source Location/Qualifiers
 1..124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="THP1"
 /note="Vector: pRIT28; Site 1: BamHI; Site 2: BamHI;
 Shotgun cloning of RDA difference products. Macrophage and
 foamcell libraries were submitted to successive rounds of
 subtractive hybridisations generating populations of gene
 fragments that are differentially expressed in macrophage
 to foam cell formation."

BASE COUNT 37 a 18 c 29 g 38 t 2 others

ORIGIN

Query Match 1.8%; Score 49; DB 10; Length 124;
 Best Local Similarity 100.0%; Pred.No. 1.8e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1551 TCGAACAGATGTGGACCGCTTTAAAGTTCATTCTTAAGCA 1599
 |||||
 76 TCGAACAGATGTGGACCGCTTTAAAGTTCATTCTTAAGCA 124

Db

RESULT 33
 BY021039 371 bp mRNA linear EST 06-DEC-2002
 LOCUS BY021039
 DEFINITION JY9-MC(A) CDNA Mus musculus gland RCB-0526
 sequence.
 ACCESSION BY021039
 VERSION BY021039.1 GI:26126482
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 371)
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
 Quackenbush, J., Schiml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
 Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,
 L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest,
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
 Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M.,
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.V., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, J.U., Reid, J., Ring,
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou,
 M., Shimada, K., Sultana, R., Takanaka, Y., Taylor, M.S., Teasdale,
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yamagisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
 Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Kono, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 K., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shidita,
 E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL MEDLINE 22354683
 PUBMED 12466851
 COMMENT Contact: Yoshinori Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```
1..371
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830044J19"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
```

BASE COUNT

```
96 a 95 c 84 g 96 t
```

Query Match

```
Best Local Similarity 1.7%; Score 47; DB 13; Length 371;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
118 TTTCAGATTGTTGAACCTTCTGCGCAGCATATACAGAGGAG 164
|||||
104 TTTCAGATTGTTGAACCTTCTGCGCAGCATATACAGAGGAG 150
```

RESULT 34

BY018262

LOCUS BY018262 401 bp mRNA linear EST 06-DEC-2002

DEFINITION BY018262 RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA Mus musculus cDNA clone G830027G18 5', mRNA

ACCESSION BY018262

VERSION BY018262.1 GI:26078511

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 401)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Miyai, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gotojori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,

TITLE

JOURNAL
MEDLINE
PUBMED

COMMENT

Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gasterland, T., Gariboldi, M., Gissi, C., Godzi, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavani, W.J., Pereira, G., Penelope, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, U., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verrardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanoisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kikukawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```
1..401
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830027G18"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
```

BASE COUNT

```
107 a 97 c 94 g 102 t 1 others
```

Query Match

```
1.7%; Score 47; DB 13; Length 401;
```

Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCACATACAGAGAGAG 164
|||||
Db 104 TTTCAGATTGTTGAACCTCTCTGCGCACATACAGAGAGAG 150

RESULT 35

BY261117

LOCUS BY261117 450 bp mRNA linear EST 10-DEC-2002

DEFINITION BY261117 RIKEN full-length enriched, visual cortex Mus musculus

ACCESSION BY261117

VERSION BY261117.1 GI:26442629

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oosato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojbori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A.,
Quackenbush, J., Schiraldi, L. M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Choitha, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarray, B. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lemhard, B., Lyons,
P. A., Maglott, D. R., Malais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pettes, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wymshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Akawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)

JOURNAL MEDLINE
PUBMED 12466851

COMMENT Contact: Yoshitake Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/

Alizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tsgami,
M., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct

Substitution
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Sequencing Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Henrich (
Laboratory for Neuronal Circuit Development Brain Science Institute
RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose
assistance we gratefully acknowledge. Please visit our web site
(http://genome.gsc.riken.go.jp) for further details.

FEATURES

source

1.450
/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="K330305G13"

/tissue_type="visual cortex"

/clone_lib="RIKEN full-length enriched, visual cortex"

BASE COUNT

130 a 96 c 102 g 122 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCACATACAGAGAGAG 164
|||||
Db 95 TTTCAGATTGTTGAACCTCTCTGCGCACATACAGAGAGAG 141

RESULT 36

BI465498

LOCUS BI465498

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE Lemishka, I., Scarce, M., Bressanelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Maira, M., Page, D., Wylie, T., Martin, J., Bistrain, A.,
Schmitt, A., Theising, B., Rittler, B., Ronko, I., Bennett, J., Cardenas,
M., Gibbons, M., McCann, K., Cole, R., Tsagaris, V., Williams, T.,
Jackson, Y. and Bowers, Y.

JOURNAL Endocrine Pancreas Consortium

COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Unpublished

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohpc.harvard.edu

Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
2000) Library was constructed by Catherine Lee DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Marie Scarce
(mscarce@mail.med.upenn.edu)

High quality sequence stop: 348.
Location/Qualifiers

FEATURES
source 1..470

GAGAGAGAGAGCGCCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGATTCGAGTTTAATTAATTAATTCCTCCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 157 a 140 c 125 g 233 t

Query Match 1.7%; Score 47; DB 10; Length 655;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 TTTCAGATTGTTGAATCTCTGCGCCGCAATATCAGAGAGAG 164
Db 298 TTTCAGATTGTTGAATCTCTGCGCCGCAATATCAGAGAGAG 344

RESULT 40
LOCUS BB658393 657 bp mRNA linear EST 26-OCT-2001
DEFINITION BB658393 RIKEN full-length enriched, 12 days embryo eyeball Mus
ACCESSION BB658393
VERSION BB658393.1 GI:16492218
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagata,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome-gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wegli,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,B., Watabe,H., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome-gsc.riken.go.jp>) for further details.
e mouse tissues.
location/Qualifiers

FEATURES
source
1. 657
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230041C01"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, 12 days embryo eyeball"
/note="Site 1: SalI; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGAGAGAGCGCCGCACTCGAGTTTCTTTTCTTTT 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGATTCGAGTTTAATTAATTAATTCCTCCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 199 a 144 c 154 g 160 t

Query Match 1.7%; Score 47; DB 10; Length 657;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 TTTCAGATTGTTGAATCTCTGCGCCGCAATATCAGAGAGAG 164
Db 318 TTTCAGATTGTTGAATCTCTGCGCCGCAATATCAGAGAGAG 364

RESULT 41
LOCUS BY734980 677 bp mRNA linear EST 17-DEC-2002
DEFINITION BY734980 RIKEN full-length enriched, mammary gland RCB-0526
ACCESSION BY734980
VERSION BY734980.1 GI:27148107
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 677)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nishida,I., Oseko,N., Saito,R., Suzuki,H., Yamataka,I., Kiyosawa,H., Yaghi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bull,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Brad,D., Brusic,V., Chochia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pereira,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sanderlin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takemata,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verrardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,

Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa M., Yang,I., Yang,L., Yuan,Z., Zavalan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,O., Aizawa,K., Aizawa,T., Fukuda,S., Hara,A., Hashizume,W., Imocani,K., Ishii,Y., Itch,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE PUBMED
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Adachi,J., Aizawa,K., Akimura,T., Aizawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imocani,K., Ishii,Y., Itch,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watanabe,A., Muramatsu,M. and Hayashizaki,Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES

source

Location/Qualifiers
1. 677
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830046C02"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Yfg-MC(A)"
/clone_1ib="RIKEN full-length enriched, mammary gland RCB-0526 Yfg-MC(A) cDNA"
BASE COUNT 130 a 167 c 149 g 224 t 7 others

Query Match 1.7%; Score 47; DB 14; Length 677;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAATCTCTGCGCCGACATACAGAGAG 164
|||||
DB 478 TTTCAGATTGTTGAATCTCTGCGCCGACATACAGAGAG 524
|||||

RESULT 42
BU708579 731 bp mRNA linear EST 26-NOV-2002
LOCUS

DEFINITION
UI-M-F10-cat-d-15-0-UI-r1 NIH BMAP F10 Mus musculus cDNA clone
IMAGE: 6416990 5', mRNA sequence.
BU708579
VERSION
BU708579.1 GI:23641181
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 731)
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bri-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLN at:
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
The following repetitive elements were found in this cDNA sequence: 35-96, >POLY_A#Simple_repeat (matched complement)
Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers
1. 731
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6416990"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1ib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Abs; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to RNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCAGCAGC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 187 a 146 c 156 g 242 t

Query Match 1.7%; Score 47; DB 13; Length 731;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAATCTCTGCGCCGACATACAGAGAG 164
|||||
DB 215 TTTCAGATTGTTGAATCTCTGCGCCGACATACAGAGAG 261
|||||

RESULT 43
CBS18391 787 bp mRNA linear EST 28-MAR-2003
LOCUS
DEFINITION
UI-M-GHO-csc-o-11-0-UI-r1 NIH BMAP_GHO Mus musculus cDNA clone
IMAGE: 6837828 5', mRNA sequence.
ACCESSION
CBS18391
VERSION
CBS18391.1 GI:29351746
KEYWORDS
EST.

SOURCE
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 787)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bhs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
The following repetitive elements were found in this cDNA
sequence: 94-126, >(GAA)n#simple_repeat (matched complement)
169-233, >POLY A#simple_repeat (matched complement)
Seq primer: pYX-5

FEATURES
source
Location/Qualifiers
1..787
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6837828"
/cissue_type="Whole Brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP GH0"
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAAGTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
Program coordinator."

BASE COUNT 200 a 160 c 170 g 256 t 1 others
ORIGIN

Query Match 1.7%; Score 47; DB 14; Length 787;
Best Local Similarity 100.0%; Pred. No. 3.5e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCGAGATTGTTGAACCTCTCTGGCCGACATACAGAGAAG-164
|||||
352 TTTCGAGATTGTTGAACCTCTCTGGCCGACATACAGAGAAG-398

Db 352 TTTCGAGATTGTTGAACCTCTCTGGCCGACATACAGAGAAG-398

RESULT 44
LOCUS AK053115 2596 bp mRNA linear HTC 05-DSC-2002
DEFINITION Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
library, clone: E030004A12 product: sema domain, immunoglobulin
domain (Ig), short basic domain, secreted, (semaphorin) 3A, full
insert sequence.
ACCESSION AK053115
VERSION AK053115.1 GI:26343226
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Carninci, P., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493774
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M.,
Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-Format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nkai, F., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombere, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohzuki, S.,
and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 665-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12108560
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imocani, K., Iehli, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takahashi, S.,
Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>
Location/Qualifiers

FEATURES

source

1.2596
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:E030004A12"
/db_xref="taxon:10090"
/clone="E030004A12"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
383..895
/note="unnamed protein product; putative
sema domain, immunoglobulin domain (Ig), short basic
domain, secreted, (semaphorin) 3A (MGI:107558,
GB|NM_009152, evidence: BLASTN, 99%, match=564)"
/codon_start=1
/protein_id="BAC35270.1"
/db_xref="GI:26343227"
/translation="MGMPGICLAFVGVLLTARANYANGNNVPRILSYKMLESN
VITNGLANSSYHTFLDEBRSRLVYGAQDHPFNLVNITKQFQITWPSVSTRDE
CKWAGDILKECANFLKVLAVAYQTHLVAGTAFAHPICLYIEVGHPRVAVICTELW
LSRGACEGTT"
2575..2580
/note="putative"
2596
/note="putative"

CDS

BASE COUNT 822 a 475 c 541 g 758 t
ORIGIN
Query Match 1.7%; Score 47; DB 11; Length 2596;
Best Local Similarity 100.0%; Pred. No. 4,8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCCACATACAGAGAG 164
|||||
Db 301 TTTCAGATTGTTGAACCTCTCTGCGCCACATACAGAGAG 347

RESULT 45
BY330392 318 bp mRNA linear EST 11-DEC-2002
LOCUS BY330392 RIKEN full-length enriched, synovial fibroblasts Mus
DEFINITION musculus cDNA clone U130014K19 5', mRNA sequence.
ACCESSION BY330392
VERSION BY330392.1 GI:26521121
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 318)
Okaaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oosato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., But, C., Hume, D. A.,
Quackenbush, J., Schriml, U. W., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chotha, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshihide Hayaishizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Nunazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watabiki, A., Muramatsu, M. and Hayaishizaki, Y. Direct
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vari, Greece) whose assistance we gratefully
acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

FEATURES

source

1..318
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="U130014K19"
/cell_type="synovial fibroblasts"
/clone_lib="RIKEN full-length enriched, synovial
fibroblasts"
BASE COUNT 92 a 72 c 75 g 79 t
ORIGIN

Query Match 1.5%; Score 41; DB 13; Length 318;
 Best Local Similarity 100.0%; Pred. No. 4,8e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GATTGTTGAACCTCTCTGCGCCGACAAATACAGACGAG 164
 DB 1 GATTGTTGAACCTCTCTGCGCCGACAAATACAGACGAG 41

RESULT 46
 CB595903 807 bp mRNA linear EST 03-APR-2003

DEFINITION AGENCOURT_13020395 NIH_MGC_178 Mus musculus cDNA clone
 IMAGE:30299607 5', mRNA sequence.

ACCESSION CB595903
 VERSION CB595903.1 GI:29513759
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 807)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Michael Brownstein Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: NDCM63 row: p column: 16
 High quality sequence stop: 471.
 Location/Qualifiers

1. 807
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30299607"
 /lab_host="DH10B (T1-Phage-resistant)"
 /clone_1ib="NIH_MGC_178"
 /note="Organ: lung and heart; Vector: pDNR-LIB; Site 1:
 SfiI (ggccattatggcc); Site 2: SfiI (ggccgctggcc); cDNA
 made by oligo-dt priming and directionally cloned. 5' and
 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACGAGTGGCATTACGCGCG-3' and
 5'-ATTCTAGAGCGCGAGCGCGCGCATG-dt(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.5 kb
 size fraction. Library created in the laboratory of M.
 Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."

BASE COUNT 260 a 175 c 204 g 168 t
 ORIGIN

Query Match 1.5%; Score 41; DB 14; Length 807;
 Best Local Similarity 100.0%; Pred. No. 6,1e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 2211
 DB 240 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 280

RESULT 47

CB196368

LOCUS

DEFINITION

IMAGE:30138167 5', mRNA sequence.

ACCESSION

CB196368

VERSION CB196368.1 GI:28224005

KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 900)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. David Rowe
 CDNA Library Preparation: Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: NDCM0040 row: 1 column: 24
 High quality sequence stop: 655.
 Location/Qualifiers

1. 900
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30138167"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_135"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Normalized full-length enriched library from pooled mouse
 embryonic 11mb, maxilla and mandible, day 12.5, 13.5, 14.5
 and 15.5 (size selected for the 0.5-1 kb fragments)
 cloned directionally, priming method: oligo-dt. cDNA
 enrichment: >1k bp. Average insert size 1.6k bp.
 Normalization (Cot value): 7.5 kb. Priming sequence:
 5'-GACTAGTCTTGAATCGGAGCGCGCGC(T)3' Tissue contributed by
 David Rowe. Library constructed by Resden, Invitrogen
 Corp."

BASE COUNT 286 a 208 c 225 g 178 t 3 others
 ORIGIN

Query Match 1.5%; Score 41; DB 14; Length 900;
 Best Local Similarity 100.0%; Pred. No. 6,3e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 2211
 DB 475 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 515

RESULT 48

B0884924

LOCUS

DEFINITION

AGENCOURT_8750939 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6336043

5', mRNA sequence.

ACCESSION B0884924

VERSION B0884924.1 GI:22276942

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 917)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Mark Macnouchie, Ph.D. and Nancy L. Freeman,
 Ph.D.
 CDNA Library Preparation: ResGen, Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.jnl.gov>
 Plate: L1AM13799 row: g column: 20
 High quality sequence stop: 724.
 Location/Qualifiers

FEATURES

source

1. .917
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6336043"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_130"
 /note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccdp;
 Site 1: BcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dt. Average insert size 1.95 kb.
 Constructed by Resgen, Invitrogen Corp. Note: this is a
 NIH MGC library."

BASE COUNT 302 a 192 c 217 g 206 t

ORIGIN

Query Match 1.5%; Score 41; DB 13; Length 917;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10; Indels 0; Gaps 0;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACCTCTTCTTAAGGTAACCTGGAAGTCATTGACACAGA 2211

DB 39 CAACCTCTTCTTAAGGTAACCTGGAAGTCATTGACACAGA 79

RESULT 49

LOCUS

CB776561 394 bp mRNA linear EST 16-MAY-2003
 AMGNNUC:SRPG2-00038-G4-A SRPG2 (10238) Rattus norvegicus cDNA clone

DERIVATION SRPG2-00038-g4 5', mRNA sequence.

VERSION CB776561 GI:29864952

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 394)

REFERENCE 1 (bases 1 to 394)

AUTHORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL

Unpublished

CONTACT: Dan Fitzpatrick

COMMENT

Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00038 row: g column: 4.

FEATURES

source

1. .394
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="srpg2-00038-g4"
 /issue_type="peneal gland brain"
 /clone_lib="srpg2 (10238)"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; peneal
 gland brain region"

BASE COUNT 127 a 88 c 78 g 101 t

ORIGIN

Query Match 1.4%; Score 38; DB 14; Length 394;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTAGATAGTCAGATATGCAAGATGCTTGGAGC 1023

DB 168 GCTAGATAGTCAGATATGCAAGATGCTTGGAGC 205

RESULT 50
 BO779621/c 657 bp mRNA linear EST 26-JUL-2002
 LOCUS UI-R-FFO-cox-f-23-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone

DEFINITION

UI-R-FFO-cox-f-23-0-UI 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 657)
 REFERENCE 1 (bases 1 to 657)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Jeff Stevens
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: DISTRIBUTION: Researchers may obtain clones
 from Research Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1. .657
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="UI-R-FFO-cox-f-23-0-UI"
 /issue_type="Mixed tissues"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-R-FFO"
 /note="Vector: pT73-Pac (Pharmacia) with a modified
 polylinker; Site_1: BcoR I; Site_2: Not I; UI-R-FFO is a
 subtracted cDNA library containing the following tissue(s)
): Normal cartilage and SR-JWS Tumor line. The
 subtraction was made according to Bonaldo, Lennon and
 Soares, Genome Research, 6:791-806, 1996. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dt)18 tail. The
 sequence tags for these libraries are: CTATGAGACG,
 CATCTGTGTA.
 TAG_LIB=UI-R-FFO
 TAG_TISSUE=cartilage
 TAG_SEQ=CTATGAGACG"

BASE COUNT

182 a 131 c 117 g 227 t

ORIGIN

Query Match 1.4%; Score 37; DB 13; Length 657;
 Best Local Similarity 100.0%; Pred. No. 8.4e-08;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2600 TTTCATGCGATTATGTGATGTTTCAATGATGGGAA 2636

DB 600 TTTCATGCGATTATGTGATGTTTCAATGATGGGAA 564

RESULT 51

BU758296/c
 LOCUS BU758296 658 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-R-FFO-cov-a-17-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone
 ACCESSION UI-R-FFO-cov-a-17-0-UI 3', mRNA sequence.
 VERSION BU758296
 KEYWORDS BU758296.1 GI:23720667
 SOURCE EST.
 ORGANISM Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 658)
 REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Jeff Stevens, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: DISTRIBUTION: Researchers may obtain clones
 from Research Genetics (www.resgen.com).
 Seq primer: MJ3 FORWARD
 POLYA=Yes

FEATURES
 source
 1..658
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="UI-R-FFO-cov-a-17-0-UI"
 /tissue_type="Mixed tissues"
 /dev_stage="Adult"
 /lab_host="UI-R-FFO"
 /clone_lib="UI-R-FFO"
 /note="Vector: pT73-Bac (Pharmacia) with a modified
 polylinker; Site 1: EcoR I; Site 2: Not I; UI-R-FFO is a
 subtracted cDNA library containing the following tissues (s
): Normal cartilage and SR-DMS Tumor line. The
 subtraction was made according to Bonaldo, Lennon and
 Soares, Genome Research, 6:791-806, 1996. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dt)18 tail. The
 sequence tags for these libraries are: CTATGACAC,
 CATCTTGTGA.
 TAG LIB=UI-R-FFO
 TAG TISSUE=cartilage
 TAG_SEQ=CTATGACAC"

BASE COUNT 181 a 131 c 117 g 227 t 2 others
 ORIGIN

Query Match 1..4*: Score 37; DB 13; Length 658;
 Best Local Similarity 100.0%; Pred No. 8.4e-08;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2600 TTTCATGGCATTATGTGATGTTTACATGGTGGAA 2636
 |||||
 DB 600 TTTCATGGCATTATGTGATGTTTACATGGTGGAA 564
 |||||

RESULT 52
 BB612039 660 bp mRNA linear EST 26-OCT-2001

DEFINITION BB612039 RIKEN full-length enriched, 15 days embryo head Mus
 musculus cDNA clone 4022435H11 5', mRNA sequence.
 ACCESSION BB612039
 VERSION BB612039.1 GI:16453118
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 660)
 REFERENCE Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
 AUTHORS Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeida,Y., Tanaka,T., Toyota,T.,
 Muramatsu,M., and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saito-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
 M., Komno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
 S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
 Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
 Y., and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanka,I., Aizawa
 K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES
 source
 1..660
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4022435H11"
 /sex="mixed"
 /tissue_type="head"
 /dev_stage="15 days embryo"
 /lab_host="DHI0B"
 /clone_lib="RIKEN full-length enriched, 15 days embryo
 head"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'

GAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCCAGTATTAATTAATATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLX I"

BASE COUNT 185 a 141 c 160 g 174 t

Query Match 1.4%; Score 37; DB 10; Length 660;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 306 ACAAGAAATGTGATCCAGATGTCATCTTT 342
DB 163 ACAAGAAATGTGATCCAGATGTCATCTTT 199

RESULT 53
LOCUS CA327319 699 bp mRNA linear EST 27-NOV-2002
DEFINITION U1-M-FY0-ccy-p-01-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6826322 5', mRNA sequence.
CA327319
ACCESSION CA327319
VERSION CA327319.1 GI:24545417
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 699)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP).
Seq primer: pyx-5.
Location/Qualifiers
1. .699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6826322"
/issue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGGAGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National

FEATURES
source

1. .699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6826322"
/issue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: Brain; Vector: pyx-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGGAGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National

BASE COUNT 197 a 165 c 156 g 180 t 1 others
Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

Query Match 1.3%; Score 35; DB 14; Length 699;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 986 GCTAGATAGTCCAGATGATGCAAGATGCTTGG 1020
DB 382 GCTAGATAGTCCAGATGATGCAAGATGCTTGG 416

RESULT 54
LOCUS BU515916 773 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10114062 NIH_MGC_134 Mus musculus cDNA clone
IMAGE: 6512603 5', mRNA sequence.
BU515916
ACCESSION BU515916
VERSION BU515916.1 GI:22823442
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 773)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLNL4084 row: p column: 12
High quality sequence stop: 616.
Location/Qualifiers
1. .773
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE: 6512603"
/issue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/note="Vector: PCMV-SPORT6.1.cdb; Site 1: EcorV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Constructed by Reegen, Invitrogen Corp. Note: this is a NIH_MGC library."

FEATURES
source
1. .773
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE: 6512603"
/issue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/note="Vector: PCMV-SPORT6.1.cdb; Site 1: EcorV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Constructed by Reegen, Invitrogen Corp. Note: this is a NIH_MGC library."

BASE COUNT 216 a 191 c 173 g 193 t
Query Match 1.3%; Score 35; DB 13; Length 773;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 986 GCTAGATAGTCCAGATGATGCAAGATGCTTGG 1020
DB 284 GCTAGATAGTCCAGATGATGCAAGATGCTTGG 318

RESULT 55
LOCUS BU703431 775 bp mRNA linear EST 09-OCT-2002
DEFINITION U1-M-FOO-bco-k-13-0-UI.r1 NIH_BMAP_FOO Mus musculus cDNA clone
IMAGE: 6405252 5', mRNA sequence.
BU703431
ACCESSION BU703431
VERSION BU703431.1 GI:23629234

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 775)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
Seq primer: PYX-5.
Location/Qualifiers
1..775
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6405252"
/issue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH BMAP F00"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 221 a 184 c 165 g 200 t 5 others

ORIGIN

Query Match 1.3%; Score 35; DB 13; Length 775;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTAGATAGTCAATATGCAAGATGACTTTGG 1020
|||||
Db 279 GCTAGATAGTCAATATGCAAGATGACTTTGG 313

RESULT 56
AA861394 374 bp mRNA linear EST 04-JAN-1999
LOCUS ak37c09.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1408144
DEFINITION 3', mRNA sequence.
ACCESSION AA861394 GI:2953534
VERSION AA861394.1 GI:2953534
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 374)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bernaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1657 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 364.
Location/Qualifiers
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1408144"
/sex="male"
/lab_host="DH10B"
/clone_1lb="Soares_testis_NHT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech laboratories Inc., and primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAAATGAGCGGCCGCCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5 and was constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 122 a 68 c 61 g 123 t

ORIGIN

Query Match 1.3%; Score 34; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GTAATTCTCAATAGGCTTTTCTTCAATATACC 2709
|||||
Db 374 GTAATTCTCAATAGGCTTTTCTTCAATATACC 341

RESULT 57
B1989239 600 bp mRNA linear EST 20-DEC-2001
LOCUS B1989239
DEFINITION 4032-91 Mouse E14.5 retina lambda ZAP II library Mus musculus cDNA, mRNA sequence.
ACCESSION B1989239 GI:17960234
VERSION B1989239.1 GI:17960234
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 600)
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Mu, X., Zhao, S., Pershad, R., Haieh, T., F., Scarpa, A., Wang, S. W., White, R. A., Beremand, P. D., Thomas, T. L., Gan, L. and Klein, W. H.
Gene expression in the developing mouse retina by EST sequencing and microarray analysis
Nucleic Acids Res. 29 (24), 4983-4993 (2001)

JOURNAL MEDLINE 21671825
PUBMED 11812828
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329.
Location/Qualifiers

source 1..600
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/clone_lib="Mouse E14.5 retina lambda ZAP II library"

BASE COUNT 197 a 134 c 142 g 127 t
ORIGIN

Query Match 1.1%; Score 30; DB 12; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2600 TTTCATGCATTATGTCATGTTTACATG 2629
Db 392 TTTCATGCATTATGTCATGTTTACATG 421

RESULT 58
LOCUS BB628362
DEFINITION BB628362 RIKEN full-length enriched, adult male urinary bladder Mus musculus cDNA clone 9530096M16 5', mRNA sequence.
ACCESSION BB628362
VERSION BB628362.1 GI:15399354
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 662)

REFERENCE
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiwamoto,K., Hori,F., Ichi,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyu,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T. et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Osawa,K., Tanaka,T., Matsumura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ichi,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, 172-186 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for

FEATURES
source
Location/Qualifiers
1..662
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="9530096M16"
/sex="male"
/tissue_type="urinary bladder"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male urinary bladder"

BASE COUNT 221 a 135 c 153 g 152 t 1 others
ORIGIN

Query Match 1.1%; Score 30; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 0.00051;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2600 TTTCATGCATTATGTCATGTTTACATG 2629
Db 270 TTTCATGCATTATGTCATGTTTACATG 299

RESULT 59
LOCUS W98303
DEFINITION w98303 690 bp mRNA linear EST 16-JUL-1996 mg13h01.r1 Soares mouse embryo NBHE13.5 14.5 Mus musculus cDNA clone IMAGE:423697 5' similar to gb:X85993 M.musculus mRNA for semaphorin D (MOUSE);, mRNA sequence.
ACCESSION W98303
VERSION W98303.1 GI:1428279
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 690)

REFERENCE
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheinberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

further details.
e mouse tissues.
Location/Qualifiers
1..662
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="9530096M16"
/sex="male"
/tissue_type="urinary bladder"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male urinary bladder"

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTGAGTTATTAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLIC I."

Db 245 CATTGGAGAGACTTCTCATTAAGATGA 273

RESULT 63

CD349263

DEFINITION

CD349263 729 bp mRNA linear EST 29-MAY-2003
UI-M-FY0-cfqa-18-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6854395 5', mRNA sequence.

ACCESSION

CD349263 GI:31140778

VERSION

EST. CD349263.1 GI:31140778

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS

1 (bases 1 to 729)
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seg primer: pyx-5.

FEATURES

Location/Qualifiers

1..729
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6854395"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5, 101.5, 102.5, 103.5, 104.5, 105.5, 106.5, 107.5, 108.5, 109.5, 110.5, 111.5, 112.5, 113.5, 114.5, 115.5, 116.5, 117.5, 118.5, 119.5, 120.5, 121.5, 122.5, 123.5, 124.5, 125.5, 126.5, 127.5, 128.5, 129.5, 130.5, 131.5, 132.5, 133.5, 134.5, 135.5, 136.5, 137.5, 138.5, 139.5, 140.5, 141.5, 142.5, 143.5, 144.5, 145.5, 146.5, 147.5, 148.5, 149.5, 150.5, 151.5, 152.5, 153.5, 154.5, 155.5, 156.5, 157.5, 158.5, 159.5, 160.5, 161.5, 162.5, 163.5, 164.5, 165.5, 166.5, 167.5, 168.5, 169.5, 170.5, 171.5, 172.5, 173.5, 174.5, 175.5, 176.5, 177.5, 178.5, 179.5, 180.5, 181.5, 182.5, 183.5, 184.5, 185.5, 186.5, 187.5, 188.5, 189.5, 190.5, 191.5, 192.5, 193.5, 194.5, 195.5, 196.5, 197.5, 198.5, 199.5, 200.5, 201.5, 202.5, 203.5, 204.5, 205.5, 206.5, 207.5, 208.5, 209.5, 210.5, 211.5, 212.5, 213.5, 214.5, 215.5, 216.5, 217.5, 218.5, 219.5, 220.5, 221.5, 222.5, 223.5, 224.5, 225.5, 226.5, 227.5, 228.5, 229.5, 230.5, 231.5, 232.5, 233.5, 234.5, 235.5, 236.5, 237.5, 238.5, 239.5, 240.5, 241.5, 242.5, 243.5, 244.5, 245.5, 246.5, 247.5, 248.5, 249.5, 250.5, 251.5, 252.5, 253.5, 254.5, 255.5, 256.5, 257.5, 258.5, 259.5, 260.5, 261.5, 262.5, 263.5, 264.5, 265.5, 266.5, 267.5, 268.5, 269.5, 270.5, 271.5, 272.5, 273.5, 274.5, 275.5, 276.5, 277.5, 278.5, 279.5, 280.5, 281.5, 282.5, 283.5, 284.5, 285.5, 286.5, 287.5, 288.5, 289.5, 290.5, 291.5, 292.5, 293.5, 294.5, 295.5, 296.5, 297.5, 298.5, 299.5, 300.5, 301.5, 302.5, 303.5, 304.5, 305.5, 306.5, 307.5, 308.5, 309.5, 310.5, 311.5, 312.5, 313.5, 314.5, 315.5, 316.5, 317.5, 318.5, 319.5, 320.5, 321.5, 322.5, 323.5, 324.5, 325.5, 326.5, 327.5, 328.5, 329.5, 330.5, 331.5, 332.5, 333.5, 334.5, 335.5, 336.5, 337.5, 338.5, 339.5, 340.5, 341.5, 342.5, 343.5, 344.5, 345.5, 346.5, 347.5, 348.5, 349.5, 350.5, 351.5, 352.5, 353.5, 354.5, 355.5, 356.5, 357.5, 358.5, 359.5, 360.5, 361.5, 362.5, 363.5, 364.5, 365.5, 366.5, 367.5, 368.5, 369.5, 370.5, 371.5, 372.5, 373.5, 374.5, 375.5, 376.5, 377.5, 378.5, 379.5, 380.5, 381.5, 382.5, 383.5, 384.5, 385.5, 386.5, 387.5, 388.5, 389.5, 390.5, 391.5, 392.5, 393.5, 394.5, 395.5, 396.5, 397.5, 398.5, 399.5, 400.5, 401.5, 402.5, 403.5, 404.5, 405.5, 406.5, 407.5, 408.5, 409.5, 410.5, 411.5, 412.5, 413.5, 414.5, 415.5, 416.5, 417.5, 418.5, 419.5, 420.5, 421.5, 422.5, 423.5, 424.5, 425.5, 426.5, 427.5, 428.5, 429.5, 430.5, 431.5, 432.5, 433.5, 434.5, 435.5, 436.5, 437.5, 438.5, 439.5, 440.5, 441.5, 442.5, 443.5, 444.5, 445.5, 446.5, 447.5, 448.5, 449.5, 450.5, 451.5, 452.5, 453.5, 454.5, 455.5, 456.5, 457.5, 458.5, 459.5, 460.5, 461.5, 462.5, 463.5, 464.5, 465.5, 466.5, 467.5, 468.5, 469.5, 470.5, 471.5, 472.5, 473.5, 474.5, 475.5, 476.5, 477.5, 478.5, 479.5, 480.5, 481.5, 482.5, 483.5, 484.5, 485.5, 486.5, 487.5, 488.5, 489.5, 490.5, 491.5, 492.5, 493.5, 494.5, 495.5, 496.5, 497.5, 498.5, 499.5, 500.5, 501.5, 502.5, 503.5, 504.5, 505.5, 506.5, 507.5, 508.5, 509.5, 510.5, 511.5, 512.5, 513.5, 514.5, 515.5, 516.5, 517.5, 518.5, 519.5, 520.5, 521.5, 522.5, 523.5, 524.5, 525.5, 526.5, 527.5, 528.5, 529.5, 530.5, 531.5, 532.5, 533.5, 534.5, 535.5, 536.5, 537.5, 538.5, 539.5, 540.5, 541.5, 542.5, 543.5, 544.5, 545.5, 546.5, 547.5, 548.5, 549.5, 550.5, 551.5, 552.5, 553.5, 554.5, 555.5, 556.5, 557.5, 558.5, 559.5, 560.5, 561.5, 562.5, 563.5, 564.5, 565.5, 566.5, 567.5, 568.5, 569.5, 570.5, 571.5, 572.5, 573.5, 574.5, 575.5, 576.5, 577.5, 578.5, 579.5, 580.5, 581.5, 582.5, 583.5, 584.5, 585.5, 586.5, 587.5, 588.5, 589.5, 590.5, 591.5, 592.5, 593.5, 594.5, 595.5, 596.5, 597.5, 598.5, 599.5, 600.5, 601.5, 602.5, 603.5, 604.5, 605.5, 606.5, 607.5, 608.5, 609.5, 610.5, 611.5, 612.5, 613.5, 614.5, 615.5, 616.5, 617.5, 618.5, 619.5, 620.5, 621.5, 622.5, 623.5, 624.5, 625.5, 626.5, 627.5, 628.5, 629.5, 630.5, 631.5, 632.5, 633.5, 634.5, 635.5, 636.5, 637.5, 638.5, 639.5, 640.5, 641.5, 642.5, 643.5, 644.5, 645.5, 646.5, 647.5, 648.5, 649.5, 650.5, 651.5, 652.5, 653.5, 654.5, 655.5, 656.5, 657.5, 658.5, 659.5, 660.5, 661.5, 662.5, 663.5, 664.5, 665.5, 666.5, 667.5, 668.5, 669.5, 670.5, 671.5, 672.5, 673.5, 674.5, 675.5, 676.5, 677.5, 678.5, 679.5, 680.5, 681.5, 682.5, 683.5, 684.5, 685.5, 686.5, 687.5, 688.5, 689.5, 690.5, 691.5, 692.5, 693.5, 694.5, 695.5, 696.5, 697.5, 698.5, 699.5, 700.5, 701.5, 702.5, 703.5, 704.5, 705.5, 706.5, 707.5, 708.5, 709.5, 710.5, 711.5, 712.5, 713.5, 714.5, 715.5, 716.5, 717.5, 718.5, 719.5, 720.5, 721.5, 722.5, 723.5, 724.5, 725.5, 726.5, 727.5, 728.5, 729.5

FEATURES

Location/Qualifiers

1..446
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G430009K02"
/clone_lib="RIKEN full-length enriched, pooled cell lines"
/note="Pooled cell lines (cell_line=RCB-0035 MEH1 164),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1, F1), (cell_line=RCB-1283 B16
melanoma), (cell_line=RCB-1702 MEH1 231
(cell_line=RCB-1734 SCA-9 clone 15), (strain=BA16/C,
(cell_type=B cells), cell_line=RCB-1669 BCL1 clone 13-20-383
(cell_type=C3H, tissue_type=brain, cell_line=RCB-1443
BC3H1)"

Query Match 1.1%; Score 29; DB 14; Length 729;
Best Local Similarity 100.0%; Pred. NO. 0.0018;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1403 TTTCGACAGATCATCAGCATGTGCA 1431

372 TTTCGACAGATCATCAGCATGTGCA 400

RESULT 64

BB859723

LOCUS

DEFINITION

BB859723 446 bp mRNA linear EST 26-NOV-2001
BB859723 RIKEN full-length enriched, pooled cell lines Mus musculus

cDNA clone G430009K02 5', mRNA sequence.

BB859723

BB859723.1 GI:17101177

VERSION

EST.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 446)

AUTHORS

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Iishi,
Y., Ito, M., Kawai, D., Kojima, Y., Komano, H., Kouda, M., Matsuyama, T.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa,
A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,
Wataniki, A., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)

JOURNAL

Unpublished

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitama-cho, Tsukuba, Ibaraki, 305-3858, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Ito,
M., Komano, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichipillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Komano, H., Fukunishi, Y., Shibata, K., Ito, M., Carninci, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

Location/Qualifiers

1..446
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G430009K02"
/clone_lib="RIKEN full-length enriched, pooled cell lines"
/note="Pooled cell lines (cell_line=RCB-0035 MEH1 164),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1, F1), (cell_line=RCB-1283 B16
melanoma), (cell_line=RCB-1702 MEH1 231
(cell_line=RCB-1734 SCA-9 clone 15), (strain=BA16/C,
(cell_type=B cells), cell_line=RCB-1669 BCL1 clone 13-20-383
(cell_type=C3H, tissue_type=brain, cell_line=RCB-1443
BC3H1)"

BASE COUNT

87 a 99 c 99 g 161 t

ORIGIN

Query Match 1.0%; Score 27; DB 10; Length 446;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGGC 144
 |||||
 377 TTTCAGATTGTTGAACCTCTCTGGC 403

RESULT 65
 AL918018 539 bp mRNA linear EST 18-SEP-2002
 LOCUS AL918018 PUR-Z1+22 Danio rerio cDNA clone 042-A05-2, mRNA sequence.
 DEFINITION AL918018
 ACCESSION AL918018
 VERSION AL918018.1 GI:23183316
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

AL918018 539 bp mRNA linear EST 18-SEP-2002
 LOCUS AL918018 PUR-Z1+22 Danio rerio cDNA clone 042-A05-2, mRNA sequence.
 DEFINITION AL918018
 ACCESSION AL918018
 VERSION AL918018.1 GI:23183316
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 539)
 Lee, S., Ruan, H., Ma, W., Xu, M., Lo, J., He, Y., Liu, F., Eun, A., Wen, Z. and Peng, J.
 15,000 unique zebrafish EST clusters from two cDNA libraries
 and Peng, J.
 Contact: Peng J
 Lab of Functional Genomics
 Institute of Molecular and Cell Biology
 30 Medical Drive, Singapore, 117609, Singapore
 Email: pengj@imcb.a-star.edu.sg
 Clone requests: pengj@imcb.a-star.edu.sg.
 Location/Qualifiers

FEATURES
 source
 1..539
 /organism="Danio rerio"
 /mol_type="mRNA"
 /strain="local wildtype"
 /db_xref="taxon:7955"
 /clone="042-A05-2"
 /tissue_type="whole embryo or fish"
 /dev_stage="mixed stages"
 /clone_1lb="PJR-Z1+22"
 /clone_1lb="PJR-Z1+22"
 139 g 133 t

BASE COUNT 141 a 126 c 139 g 133 t

ORIGIN
 Query Match 1.0%; Score 27; DB 9; Length 539;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 GATGAATGCAGTGGCTGGAAAGAC 526
 |||||
 250 GATGAATGCAGTGGCTGGAAAGAC 276

RESULT 66
 BG926854 685 bp mRNA linear EST 06-NOV-2001
 LOCUS BG926854 HNC20-1-B9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
 DEFINITION BG926854
 ACCESSION BG926854
 VERSION BG926854.1 GI:14321377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Human sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 685)
 Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sathe, G., Mui, P., Agarwal, P., Badger, A.M., Lee, D.C., Gowen, M. and Laik, M.W.
 Identification and initial characterization of 5000 expressed

REFERENCE
 AUTHORS Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J., Sathe, G., Mui, P., Agarwal, P., Badger, A.M., Lee, D.C., Gowen, M. and Laik, M.W.
 Identification and initial characterization of 5000 expressed

sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries
 Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 PUBMED 11597177
 COMMENT Contact: Sanjay Kumar
 UW2109
 GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-1@gsk.com
 Seq primer: T7.
 Location/Qualifiers

FEATURES
 source
 1..685
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="Cartilage"
 /lab_host="E.coli DH10 B"
 /clone_1lb="HNC (Human Normal Cartilage)"
 /note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI; Directional"

BASE COUNT 178 a 150 c 193 g 164 t

ORIGIN
 Query Match 0.9%; Score 25; DB 12; Length 685;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGAGTAGG 81
 |||||
 77 GTGACCCACGCGTCGGAGTAGG 101

RESULT 67
 CD203422 556 bp mRNA linear EST 21-MAY-2003
 LOCUS CD203422 Ls AM1_08D06 T7 Litomosoides sigmodontis adult male 1 Litomosoides
 DEFINITION Ls AM1_08D06 T7 Litomosoides sigmodontis adult male 1 Litomosoides
 sigmodontis cDNA clone Ls AM1_08D06 5' similar to P90699 (P90699)
 Putative RNA binding protein, mRNA sequence.
 CD203422
 CD203422.1 GI:30931437
 EST.
 VERSION CD203422.1 GI:30931437
 KEYWORDS Litomosoides sigmodontis
 SOURCE Litomosoides sigmodontis
 ORGANISM Litomosoides sigmodontis

Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Litomosoides.
 1 (bases 1 to 556)
 LeGeoff, J., Thomas, F., Allen, J. and Blaxter, M.L.
 Genes expressed in adult male Litomosoides sigmodontis
 Unpublished
 Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JT, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was constructed by laetitia LeGeoff in Judith Allen's
 laboratory at ICAPB, University of Edinburgh. Sequencing performed
 by Fran Thomas, in Mark Blaxter's laboratory, ICAPB, University of
 Edinburgh. The sequence contained a PolyA tail (trimmed)

PCR Primers
 FORWARD: M13L
 BACKWARD: M13R
 Place: 08 row: D column: 06
 Seq primer: T7
 High quality sequence start: 8
 High quality sequence stop: 533.
 Location/Qualifiers

FEATURES
 source
 1..556

```

/organism="Litomosoides sigmodontis"
/mol_type="mRNA"
/db_xref="taxon:42156"
/clone="Ls_AM1_08006"
/sex="male"
/dev_stage="adult"
/clone_lib="Litomosoides sigmodontis adult male 1"
/notes="Vector: pSPORT-1; Site 1: Sal I; Site 2: Not I; The library was constructed by Laetitia Legeoff in Judith Allen's laboratory at ICAPB, University of Edinburgh. The library is directional, with the 5' end of inserts adjacent to the Sal I site; sequencing with the T7 primer thus gives 5' ESTs."
BASE COUNT      112 a      114 c      148 g      182 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      57 GTCGACCCACGCGTCCGGAGTAG 80
|||||
Db      12 GTCGACCCACGCGTCCGGAGTAG 35

RESULT 68      598 bp      mRNA      linear      EST 21-MAY-2003
CD203073      Ls_AM1_02E03_T7 Litomosoides sigmodontis adult male 1 Litomosoides
LOCUS      sigmodontis cDNA clone Ls_AM1_02E03 5' similar to MSP2_ONCVO
ACCESSION      (P13263) Major sperm protein 2 (MSP2), mRNA sequence.
VERSION      CD203073
KEYWORDS      CD203073.1 GI:30930640
SOURCE      EST.
ORGANISM      Litomosoides sigmodontis
LITOMOSOIDES sigmodontis
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Litomosoides.
1 (bases 1 to 598)
REFERENCE      Leggeoff, L., Thomas, F., Allen, J. and Blaxter, M.L.
AUTHORS      Genes expressed in adult male Litomosoides sigmodontis
TITLE      Unpublished
JOURNAL      Contact: Blaxter ML
COMMENT      Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was constructed by Laetitia Legeoff in Judith Allen's
laboratory at ICAPB, University of Edinburgh. Sequencing performed
by Fran Thomas, in Mark Blaxter's laboratory, ICAPB, University of
Edinburgh. The sequence contained a PolyA tail (trimmed)
PCR Primers
FORWARD: M13L
BACKWARD: M13R
Plate: 02 row: E column: 03
Seq primer: T7
High quality sequence start: 12
High quality sequence stop: 531.
Location/Qualifiers
1..598
/organism="Litomosoides sigmodontis"
/mol_type="mRNA"
/db_xref="taxon:42156"
/clone="Ls_AM1_02E03"
/sex="male"
/dev_stage="adult"
/clone_lib="Litomosoides sigmodontis adult male 1"
/notes="Vector: pSPORT-1; Site 1: Sal I; Site 2: Not I; The library was constructed by Laetitia Legeoff in Judith Allen's laboratory at ICAPB, University of Edinburgh. The

```

```

library is directional, with the 5' end of inserts
adjacent to the Sal I site; sequencing with the T7 primer
thus gives 5' ESTs."
BASE COUNT      156 a      151 c      157 g      134 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 598;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      57 GTCGACCCACGCGTCCGGAGTAG 80
|||||
Db      32 GTCGACCCACGCGTCCGGAGTAG 55

RESULT 69      614 bp      mRNA      linear      EST 14-NOV-2002
CA502146      WHE4043_C11_P21ZT wheat meiotic anther cDNA library Triticum
LOCUS      aestivum cDNA clone WHE4043_C11_P21, mRNA sequence.
DEFINITION      CA502146
ACCESSION      CA502146
KEYWORDS      CA502146.1 GI:24993106
SOURCE      EST.
ORGANISM      Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 614)
REFERENCE      Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R., Pham
AUTHORS      J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
TITLE      The structure and function of the expressed portion of the wheat
JOURNAL      genomes - Meiotic anther cDNA library
COMMENT      Unpublished
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: T7 primer.
Location/Qualifiers
1..614
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4043_C11_P21"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="wheat meiotic anther cDNA library"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
florets. One anther was sacrificed for microscopic staging
and if determined to be between (and including) meiotic
stages pre-meiosis and metaphase I, the remaining two
anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge Lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT      168 a      164 c      163 g      119 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 614;

```

Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAGTAG 81
|||||
Db 1 TCGACCCACGCGTCCGGAGTAG 24

RESULT 70

LOCUS BG781873 792 bp mRNA linear EST 20-MAY-2001
DEFINITION Strongylocentrotus purpuratus cDNA clone PC_0013_B1_B03_MR_5', mRNA sequence.

ACCESSION

VERSION BG781873
KEYWORDS BG781873.1 GI:14152886

SOURCE

ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidae; Euechinoidea; Echinacea; Echinoidae;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 792)
Zhu,X., Mahaffey,G., Illies,M.R., Cameron,R.A., Davidson,E.H. and
Ettensohn,C.A.

TITLE A large scale analysis of mRNAs expressed by primary mesenchyme
cells of the sea urchin embryo

JOURNAL Development 128 (13), 2615-2627 (2001)
MEDLINE 21384984
PubMed 11493577

COMMENT Contact: Ettensohn CA
Dept. Biol. Sci.
Carnegie Mellon University
4400 Fifth Avenue, Pittsburgh, PA 15213, USA
Tel: +1 412 268 5849
Email: ettensohn@andrew.cmu.edu.

FEATURES

source

1..792
/organism="Strongylocentrotus purpuratus"

/mol_type="mRNA"

/db_xref="taxon:7668"

/clone="PC_0013_B1_B03_MR"

/tissue_type="embryo"

/cell_type="primary mesenchyme cells"

/lab_host="E.coli"

/clone_lib="Sea urchin primary mesenchyme cell cDNA
library"

/note="Vector: pSPORT1, Site_1: NotI; Site_2: SalI; oligo
dt priming from poly A+ RNA, directionally cloned"

BASE COUNT 227 a 233 c 188 g 144 t

ORIGIN

Query Match 0.9%; Score 24; DB 12; Length 792;
Best Local Similarity 100.0%; Pred. No. 0.94;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGTAG 80
|||||
Db 20 GTGACCCACGCGTCCGGAGTAG 43

RESULT 71

LOCUS BU255423 825 bp mRNA linear EST 26-NOV-2002
DEFINITION 603414282F1 CSEQCHN38 Gallus gallus cDNA clone CHEST336c22 5', mRNA sequence.

ACCESSION BU255423
VERSION BU255423.1 GI:25512564

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE

Phasianinae; Gallus.
1 (bases 1 to 825)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..825
/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST336c22"

/dev_stage="36"

/lab_host="DH10B"

/clone_lib="CSEQCHN38"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adaptors, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996) 791, except that a significantly longer
reannealing hybridization was used."

BASE COUNT 288 a 151 c 149 g 236 t 1 others
ORIGIN

Query Match 0.9%; Score 24; DB 13; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.95;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2578 AATGCAATATACATGACTTTT 2601
|||||
Db 151 AATGCAATATACATGACTTTT 174

RESULT 72

LOCUS BU326664 851 bp mRNA linear EST 28-NOV-2002
DEFINITION 603493391F1 CSEQCHN63 Gallus gallus cDNA clone CHEST339e23 5', mRNA sequence.

ACCESSION BU326664
VERSION BU326664.1 GI:25834665

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 851)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

```
1. 851
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHST80711"
/dev_stage="22"
/lab_host="DH10B"
/clone_1ib="CSECHN52"
/notes="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
```

BASE COUNT
ORIGIN

295 a 153 c 154 g 249 t

Query Match 0.9%; Score 24; DB 13; Length 851;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2578 AATGCAATATACATGAACTTTT 2601

128 AATGCAATATACATGAACTTTT 151

RESULT 73
LOCUS BU267223 897 bp mRNA linear EST 26-NOV-2002
DEFINITION 603815106r1 CSECHN52 Gallus gallus cDNA clone CHEST80711 5', mRNA
ACCESSION BU267223
VERSION BU267223.1 GI:25538173
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 897)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
12445392
PUBMED
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1. 897

```
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHST80711"
/dev_stage="22"
/lab_host="DH10B"
/clone_1ib="CSECHN52"
/notes="Organ: limbs; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
```

BASE COUNT
ORIGIN

306 a 172 c 174 g 245 t

Query Match 0.9%; Score 24; DB 13; Length 897;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2578 AATGCAATATACATGAACTTTT 2601

192 AATGCAATATACATGAACTTTT 215

RESULT 74
LOCUS AM080828/c 184 bp mRNA linear EST 14-OCT-1999
DEFINITION xc38a05.x1 NCI CGAP Co20 Homo sapiens cDNA clone IMAGE:2586512 3',
similar to gb:J02765_ccl1 CALCYCLIN (HUMAN);, mRNA sequence.
ACCESSION AM080828
VERSION AM080828.1 GI:6035980
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 184)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES

source

```
1. 184
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2586512"
/tissue_type="moderately differentiated adenocarcinoma"
/lab_host="DH10B"
/clone_1ib="NCI CGAP Co20"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;  
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
```

Normalized to Cot 500. Average insert size 1.1kb.
 Normalized version of NCI_CGAP_CO18. Library constructed
 by Life Technologies."

BASE COUNT 50 a 46 c 38 g 50 t

ORIGIN

Query Match 0.8%; Score 23; DB 9; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAGTA 79
 ||||||||||||||||||
 DB 179 GTGACCCACGCGTCCGGAGTA 157

RESULT 75

BG898214

LOCUS 401 bp mRNA linear EST 06-NOV-2001

DEFINITION HOA38-1-B10 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA,

mRNA sequence.

ACCESSION BG898214 GI:14308463

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED
 21482651
 11597177

COMMENT
 Contact: Sanjay Kumar
 UM2109

GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-1@sk.com

Seq primer: T7.

FEATURES
 Location/Qualifiers

1..401

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="cartilage"

/lab_host="E. coli DH10 B"

/clone_lib="HOA (Human Osteoarthritic Cartilage)"

/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
 Directional"

BASE COUNT 67 a 132 c 130 g 72 t

ORIGIN

Query Match 0.8%; Score 23; DB 12; Length 401;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAGTA 79
 ||||||||||||||||||
 DB 69 GTGACCCACGCGTCCGGAGTA 91

Search completed: July 31, 2003, 23:52:32
 Job time : 5037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 17:17:31 ; Search time 679 Seconds

(without alignments)
10769.923 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aactcttattatcatgatg.....agcttttttctcatatacc 2709

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2709	100.0	2709	22	AAH47049
2	2508	92.6	2530	19	AAK75767
3	2508	92.6	2530	24	AAK75767
4	2508	92.6	2601	16	AAO87442
5	1253	46.3	1481	16	AAO29331
6	405	15.0	456	23	ABK45213
7	405	15.0	456	24	ABK45213
8	303	11.2	354	22	AAI29182

9	303	11.2	354	25	ABK23368	Human colon tumour
10	292	10.8	496	23	ABK52213	Human liver single
11	292	10.8	496	24	ABK52213	Human genome-deriv
12	160	5.9	172	22	ABA71205	Human foetal liver
13	160	5.9	172	22	AAK19504	Human brain expres
14	160	5.9	172	22	AAK45495	Human bone marrow
15	160	5.9	172	22	AAI51439	Probe #20125 used
16	160	5.9	172	22	ABK45495	Human liver single
17	160	5.9	172	22	ABK45495	Human genome-deriv
18	160	5.9	172	22	ABK45495	Human foetal liver
19	160	5.9	172	22	ABK45495	Human brain expres
20	160	5.9	172	22	ABK45495	Human bone marrow
21	160	5.9	172	22	ABK45495	Probe #20125 used
22	160	5.9	172	22	ABK45495	Human liver single
23	160	5.9	172	22	ABK45495	Human genome-deriv
24	160	5.9	172	22	ABK45495	Human foetal liver
25	160	5.9	172	22	ABK45495	Human brain expres
26	160	5.9	172	22	ABK45495	Human bone marrow
27	160	5.9	172	22	ABK45495	Probe #20125 used
28	160	5.9	172	22	ABK45495	Human liver single
29	160	5.9	172	22	ABK45495	Human genome-deriv
30	160	5.9	172	22	ABK45495	Human foetal liver
31	160	5.9	172	22	ABK45495	Human brain expres
32	160	5.9	172	22	ABK45495	Human bone marrow
33	160	5.9	172	22	ABK45495	Probe #20125 used
34	160	5.9	172	22	ABK45495	Human liver single
35	160	5.9	172	22	ABK45495	Human genome-deriv
36	160	5.9	172	22	ABK45495	Human foetal liver
37	160	5.9	172	22	ABK45495	Human brain expres
38	160	5.9	172	22	ABK45495	Human bone marrow
39	160	5.9	172	22	ABK45495	Probe #20125 used
40	160	5.9	172	22	ABK45495	Human liver single
41	160	5.9	172	22	ABK45495	Human genome-deriv
42	160	5.9	172	22	ABK45495	Human foetal liver
43	160	5.9	172	22	ABK45495	Human brain expres
44	160	5.9	172	22	ABK45495	Human bone marrow
45	160	5.9	172	22	ABK45495	Probe #20125 used
46	160	5.9	172	22	ABK45495	Human liver single
47	160	5.9	172	22	ABK45495	Human genome-deriv
48	160	5.9	172	22	ABK45495	Human foetal liver
49	160	5.9	172	22	ABK45495	Human brain expres
50	160	5.9	172	22	ABK45495	Human bone marrow
51	160	5.9	172	22	ABK45495	Probe #20125 used
52	160	5.9	172	22	ABK45495	Human liver single
53	160	5.9	172	22	ABK45495	Human genome-deriv
54	160	5.9	172	22	ABK45495	Human foetal liver
55	160	5.9	172	22	ABK45495	Human brain expres
56	160	5.9	172	22	ABK45495	Human bone marrow
57	160	5.9	172	22	ABK45495	Probe #20125 used
58	160	5.9	172	22	ABK45495	Human liver single
59	160	5.9	172	22	ABK45495	Human genome-deriv
60	160	5.9	172	22	ABK45495	Human foetal liver
61	160	5.9	172	22	ABK45495	Human brain expres
62	160	5.9	172	22	ABK45495	Human bone marrow
63	160	5.9	172	22	ABK45495	Probe #20125 used
64	160	5.9	172	22	ABK45495	Human liver single
65	160	5.9	172	22	ABK45495	Human genome-deriv
66	160	5.9	172	22	ABK45495	Human foetal liver
67	160	5.9	172	22	ABK45495	Human brain expres
68	160	5.9	172	22	ABK45495	Human bone marrow
69	160	5.9	172	22	ABK45495	Probe #20125 used
70	160	5.9	172	22	ABK45495	Human liver single
71	160	5.9	172	22	ABK45495	Human genome-deriv
72	160	5.9	172	22	ABK45495	Human foetal liver
73	160	5.9	172	22	ABK45495	Human brain expres
74	160	5.9	172	22	ABK45495	Human bone marrow
75	160	5.9	172	22	ABK45495	Probe #20125 used
76	160	5.9	172	22	ABK45495	Human liver single
77	160	5.9	172	22	ABK45495	Human genome-deriv
78	160	5.9	172	22	ABK45495	Human foetal liver
79	160	5.9	172	22	ABK45495	Human brain expres
80	160	5.9	172	22	ABK45495	Human bone marrow
81	160	5.9	172	22	ABK45495	Probe #20125 used


```
C 82 22 0.8 9072 20 AAX07558 Synthetic modified
C 83 22 0.8 9183 24 AAD28270 Alpha-lactalbumin
C 84 22 0.8 9183 24 AAD28309 Alpha-lactalbumin
C 85 21 0.8 29 25 ABZ25420 Human semaphorin-3
C 86 21 0.8 332 25 ABX61182 Arabidopsis thaliana
C 87 21 0.8 341 22 AAL19287 Human breast cancer
C 88 21 0.8 355 22 ABA08285 Human protein-Tyr
C 89 21 0.8 442 21 AAF16393 Human prostate can
C 90 21 0.8 457 25 ABX39208 Bovine EST associa
C 91 21 0.8 489 22 AAK57109 Human immune/haema
C 92 21 0.8 578 25 ABZ53641 Aspergillus oryzae
C 93 21 0.8 717 22 AAK34766 Human colon cancer
C 94 21 0.8 730 24 ABS77215 Frog embryonic gen
C 95 21 0.8 745 24 ABS77448 Frog embryonic gen
C 96 21 0.8 749 24 ABS77126 Frog embryonic gen
C 97 21 0.8 749 24 ABS77195 Frog embryonic gen
C 98 21 0.8 750 24 ABS77132 Frog embryonic gen
C 99 21 0.8 750 24 ABS77271 Frog embryonic gen
C 100 21 0.8 781 24 ABS77150 Frog embryonic gen
```

ALIGNMENTS

```
RESULT 1
AAH47049 standard; DNA; 2709 BP.
AAH47049;
29-OCT-2001 (first entry)
Semaphorin D cDNA sequence.
DE
XX Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;
XX 24p; proliferin; maspin; cancer; cytostatic; gene therapy; ss.
XX Homo sapiens.
OS
XX WO200155455-A2.
XX 02-AUG-2001.
XX 31-JAN-2001; 2001WO-US03161.
XX 31-JAN-2000; 2000US-0179191.
XX (MILL-) MILLENNIUM PHARM INC.
XX (JINS/) JIN S.
XX Jin S;
XX WPI: 2001-488799/53.
XX
XX Determining if a compound modulates the drug resistance of a cell,
XX comprises determining the expression or activity level of a resistance
XX sequence in a cell in the presence of the test compound -
XX
XX Example 1; Fig 1A-B; 79pp; English.
XX
XX The invention relates to a method of determining whether a test compound
XX modulates the drug resistance of a cell that comprises determining the
XX expression or activity level of resistance genes (e.g. semaphorin D, B94,
XX mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
XX the test compound, and comparing its expression or activity level in a
XX cell without the test compound. The drug resistant sequences are useful
XX in identifying drug resistant cells, in screening methods directed to the
XX identification of compounds that can modulate the drug resistance of a
XX cell type or multiple cell types. An isolated resistance protein can be
XX used as an immunogen to generate antibodies that bind the resistance
XX protein. Resistance nucleic acids may be inserted into vectors and used
XX as gene therapy vectors. An anti-resistance protein antibody may be used
XX to isolate a resistance protein, or facilitate the purification of
```

CC natural resistance protein from cells and of recombinantly produced
CC resistance protein expressed in host cells. The methods are useful for
CC treating a subject having a disorder, such as a drug-resistance cancer,
CC characterized by aberrant resistance sequence expression or activity by
CC administering to the subject a resistance modulator. The present sequence
CC represents a semaphorin cDNA sequence, whose expression was increased in
CC drug resistant EMT6 tumours.
XX

SQ Sequence 2709 BP; 831 A; 560 C; 620 G; 698 T; 0 other;

Query Match 100.0%; Score 2709; DB 22; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AATCTTATTTATTCAGATGTTAACAACCTAGTAATGATGCCACCTGAGGGGTGCG 60
DB 1 AATCTTATTTATTCAGATGTTAACAACCTAGTAATGATGCCACCTGAGGGGTGCG 60
QY 61 ACCCAGCGGTCGGGAGTAGGTGAGCTCGGCTGTTCCCATTTGACGCCAGTCTATT 120
DB 61 ACCCAGCGGTCGGGAGTAGGTGAGCTCGGCTGTTCCCATTTGACGCCAGTCTATT 120
QY 121 CCAGATTTGTTGAACCTCTCTGGCCGCAATATACAGAGAGAGACTAAAGCAAGG 180
DB 121 CCAGATTTGTTGAACCTCTCTGGCCGCAATATACAGAGAGAGACTAAAGCAAGG 180
QY 181 GACCTACAGCGCTCGACGACATGAGCTGTTAAGTATGATGACCACTGAGGGGTGCG 240
DB 181 GACCTACAGCGCTCGACGACATGAGCTGTTAAGTATGATGACCACTGAGGGGTGCG 240
QY 241 ATTACTTACAGCAAGCAAACTATCAGATGGAGAGAAACATGTGCCAAGGCTGAATT 300
DB 241 ATTACTTACAGCAAGCAAACTATCAGATGGAGAGAAACATGTGCCAAGGCTGAATT 300
QY 301 ATCTCAACAAGAAATGTTGGATCCAAATGATGATCATCTTCAATGAGCTTGCCAA 360
DB 301 ATCTCAACAAGAAATGTTGGATCCAAATGATGATCATCTTCAATGAGCTTGCCAA 360
QY 361 CTCGAGTATCATACCTCTCTTGGATGAGAAACGAGTAGGCTGATGTTGAGAGAA 420
DB 361 CTCGAGTATCATACCTCTCTTGGATGAGAAACGAGTAGGCTGATGTTGAGAGAA 420
QY 421 GGATCACAATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTGTGGCC 480
DB 421 GGATCACAATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTGTGGCC 480
QY 481 AGTATCTTACACCGAAGAGATGATCAAGTGGCTGGAAGAACATCTCTGAAGAAATG 540
DB 481 AGTATCTTACACCGAAGAGATGATCAAGTGGCTGGAAGAACATCTCTGAAGAAATG 540
QY 541 TGCTAATTTTCAAGGTAAGGATATATCAGATCACTGTAAGCGCTGTGAGAC 600
DB 541 TGCTAATTTTCAAGGTAAGGATATATCAGATCACTGTAAGCGCTGTGAGAC 600
QY 601 GGGGGCTTTTTCATCCATTTTGACCTTACATTTGAATTTGACATCATCTGAGACAATAT 660
DB 601 GGGGGCTTTTTCATCCATTTTGACCTTACATTTGAATTTGACATCATCTGAGACAATAT 660
QY 661 TTTTAACTGGAGAACTCAATTTTGAAGACGCGCTGGAGAGTCCATATGACCTTAA 720
DB 661 TTTTAACTGGAGAACTCAATTTTGAAGACGCGCTGGAGAGTCCATATGACCTTAA 720
QY 721 GCTGCTACACGATCCCTTTTAATAGTAGAATTAATCTGGAACCTGAGCTGATTT 780
DB 721 GCTGCTACACGATCCCTTTTAATAGTAGAATTAATCTGGAACCTGAGCTGATTT 780
QY 781 TATGGGCGAGACTTTGCTATCTTCCGAACTTTGGGACACACCAATCAGACAGA 840
DB 781 TATGGGCGAGACTTTGCTATCTTCCGAACTTTGGGACACACCAATCAGACAGA 840
QY 841 GCAGCATGATTTCCAGGTGCTCAATGATCCAAAGTTATTAATGAGTCCCATCTTCA 900
DB 841 GCAGCATGATTTCCAGGTGCTCAATGATCCAAAGTTATTAATGAGTCCCATCTTCA 900
```

OY	901	GAGTGCATTCCTGGAAGAAGACAAAGATACTTTTCTCCCGGAAAATGCCATPACATGG	960
Db	901	GAGTGCAATTCCTGGAAGATGACAAAGTATCTTTTCTTCGGAAAAATGCATPAGTGG	960
OY	961	AGAACACTCTGGAAAAAGCTACTCAGCCGTAGAATAGGTCAGATATGCAAGATGACTTTGG	102
Db	961	AGAACACTCTGGAAAAAGCTACTCAGCCGTAGAATAGGTCAGATATGCAAGATGACTTTGG	102
OY	1021	AGGGCACAGAACTCTGGTGTATTAATTGSCAACATTCCTCAAGCTCGTCTGATTTGGCTC	108
Db	1021	AGGGCACAGAACTCTGGTGTATTAATTGSCAACATTCCTCAAGCTCGTCTGATTTGGCTC	108
OY	1081	AGTGCAGGTCCAAATNGCATTTGACACACTATTTTGTATGAACTGCAGAGATGTATTCCTAAT	114
Db	1081	AGTGCAGGTCCAAATNGCATTTGACACACTATTTTGTATGAACTGCAGAGATGTATTCCTAAT	114
OY	1141	GAACTTTAAAGATCCTAAAANAATCCAGTTGTATATGGAAGTGTATTCAGACTTCCAGTAACAT	120
Db	1141	GAACTTTAAAGATCCTAAAANAATCCAGTTGTATATGGAAGTGTATTCAGACTTCCAGTAACAT	120
OY	1201	TTTTCAAGGGATACAGCCGTGTGTATGTATAGCATAGATGTATGTGAAGAAAGGTGTCTTGG	126
Db	1201	TTTTCAAGGGATACAGCCGTGTGTATGTATAGCATAGATGTATGTGAAGAAAGGTGTCTTGG	126
OY	1261	TCCATATGCCCCACAGGGATGAGCCCACTATCAATGGGTGCTTATCAAGSAAAGATCCC	132
Db	1261	TCCATATGCCCCACAGGGATGAGCCCACTATCAATGGGTGCTTATCAAGSAAAGATCCC	132
OY	1321	CTATCCAGGGCCAGGACCTTGTCCAGCAAAAATTTGGTGTTTGACTTCACAAAGAA	138
Db	1321	CTATCCAGGGCCAGGACCTTGTCCAGCAAAAATTTGGTGTTTGACTTCACAAAGAA	138
OY	1381	CCTTCCTATGATGTATTAACCTTTGCAAGAAAGTACCCAGGCATGTACAATCCAGTGT	144
Db	1381	CCTTCCTATGATGTATTAACCTTTGCAAGAAAGTACCCAGGCATGTACAATCCAGTGT	144
OY	1441	TCTATGAAACAATCGCCCCAATAGTATCAAAAACGATGTAAATTATCAATTTTACACAAT	150
Db	1441	TCTATGAAACAATCGCCCCAATAGTATCAAAAACGATGTAAATTATCAATTTTACACAAT	150
OY	1501	TGTGCTAACCAGAGTGTATGCAAGAAATGAGACAGATATGTTATGTATTTATCGGAACAG	156
Db	1501	TGTGCTAACCAGAGTGTATGCAAGAAATGAGACAGATATGTTATGTATTTATCGGAACAG	156
OY	1561	TGTTGGGACCGTTCTTAAAGTAGTTCATTTCTTAAGAGACCTTGGTATGATTTAGAAAG	162
Db	1561	TGTTGGGACCGTTCTTAAAGTAGTTCATTTCTTAAGAGACCTTGGTATGATTTAGAAAG	162
OY	1621	GGTTCTGCTGGAAGAAATGACAGTTTTCGGSAACCGACTGTATTCAGCAATGAGGCT	168
Db	1621	GGTTCTGCTGGAAGAAATGACAGTTTTCGGSAACCGACTGTATTCAGCAATGAGGCT	168
OY	1681	TTCCACTAAGCGCAACAACCTATATATGGTTCAACGCGTGGGGTTGCCACACTCCCTT	174
Db	1681	TTCCACTAAGCGCAACAACCTATATATGGTTCAACGCGTGGGGTTGCCACACTCCCTT	174
OY	1741	AACCCGGTGTATTTTACOGGAAAGCGTGTCTGAGTGTGCTCGCCCGAGACCCCTTA	180
Db	1741	AACCCGGTGTATTTTACOGGAAAGCGTGTCTGAGTGTGCTCGCCCGAGACCCCTTA	180
OY	1801	CTGTGCTTGGAGTGGTTCTGCATGTTCCTCGCTATTTTCCCATCGCAAMAGACGACAAAG	186
Db	1801	CTGTGCTTGGAGTGGTTCTGCATGTTCCTCGCTATTTTCCCATCGCAAMAGACGACAAAG	186
OY	1861	ACGACAAAGATATTAAGAAATGAGACCCACTGACTACGTGTCAGACTTACACATGATTA	192
Db	1861	ACGACAAAGATATTAAGAAATGAGACCCACTGACTACGTGTCAGACTTACACATGATTA	192
OY	1921	TCAACATGGCCACAGCCCTGAAAGAGAAATCATCTATGGTGTAGAAGATATGACACATT	198
Db	1921	TCAACATGGCCACAGCCCTGAAAGAGAAATCATCTATGGTGTAGAAGATATGACACATT	198

QY		1981	TTTGGAAATGCACTCCGAAGTGCGACAGAGCCCTGGTCTATTGGCAATTCCAGAGCCGAAA	2040
Db		1981	TTTTGAATGCAGTCCGAAGTCGACAGAGCCCTGGTCTATTGGCAATTCCAGAGCCGAAA	2040
QY		2041	TGAAGACGGAAAAAGAAGATCAGAGTGATGATCATTCATCAAGGACAGATCAAGGCCCT	2100
Db		2041	TGAAGACGGAAAAAGAAGATCAGAGTGATGATCATTCATCAAGGACAGATCAAGGCCCT	2100
QY		2101	TCGTCTACGTAGCTTACACAAGAGGATTCAGGCAATTACTCTGCCATGCGGTGGAACA	2160
Db		2101	TCGTCTACGTAGCTTACACAAGAGGATTCAGGCAATTACTCTGCCATGCGGTGGAACA	2160
QY		2161	TGGGTTCTATCAAACCTCTTCTTAAGAGTACCCTGGAAGTCAATTGACACAGACATTTGGA	2220
Db		2161	TGGGTTCTATCAAACCTCTTCTTAAGAGTACCCTGGAAGTCAATTGACACAGACATTTGGA	2220
QY		2221	AGAACTTCTTCATPAAAGATGATGATGGAGATGGCTTAAAGCCAAAGAAATGTCCTAATG	2280
Db		2221	AGAACTTCTTCATPAAAGATGATGATGGAGATGGCTTAAAGCCAAAGAAATGTCCTAATG	2280
QY		2281	CATGACACCTTAGCCAGAAAGGTCTGTGTACAGACCTTCATGACGCTCATCAACCCCAA	2340
Db		2281	CATGACACCTTAGCCAGAAAGGTCTGTGTACAGACCTTCATGACGCTCATCAACCCCAA	2340
QY		2341	TCTCAACACGATGGATGACTTCTGTGAACAA GTTTGAAAGGACCGAAACAACTCG	2400
Db		2341	TCTCAACACGATGGATGACTTCTGTGAACAA GTTTGAAAGGACCGAAACAACTCG	2400
QY		2401	GCAAAAGCCAGACATPACCCTCCAGGGAA CAGTAA CAATNGAAGCCTTACAGAAATPA	2460
Db		2401	GCAAAAGCCAGACATPACCCTCCAGGGAA CAGTAA CAATNGAAGCCTTACAGAAATPA	2460
QY		2461	GAAAGGTAGAAA CAGAGAACCCACGA TTTGAGAGGGCACCCAGAGTGTCTGAGCTGC	2520
Db		2461	GAAAGGTAGAAA CAGAGAACCCACGA TTTGAGAGGGCACCCAGAGTGTCTGAGCTGC	2520
QY		2521	ATTACCTCTAGAAACCTCAAA CAGTGAATCTTGCTTGAACATAATCTGAAAAACAA	2580
Db		2521	ATTACCTCTAGAAACCTCAAA CAGTGAATCTTGCTTGAACATAATCTGAAAAACAA	2580
QY		2581	TGCAATATPACATGAACTTTTTCATGSCATTAATGATGATTTTCAATGGTGGAAATTC	2640
Db		2581	TGCAATATPACATGAACTTTTTCATGSCATTAATGATGATTTTCAATGGTGGAAATTC	2640
QY		2641	AGCTGAGTCCACCAATTAATTAATTAATCATGAGTAACTTTTCTTAATAGGCTTTT	2700
Db		2641	AGCTGAGTCCACCAATTAATTAATTAATCATGAGTAACTTTTCTTAATAGGCTTTT	2700
QY		2701	CCCTAATACC 2709	
Db		2701	CCCTAATACC 2709	
RESULT 2				
ID	AAAX75767		standard; DNA; 2530 BP.	
AC	AAAX75767;			
XX				
DT	22-JUL-1999	(first entry)		
DE				
XX				
XX				
KW	Human; beta-amyloid precursor protein; beta-AP; diagnosis; cancer; framehift mutation; age-related disease; neurodegenerative disorder; Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; Huntington's disease; multiple sclerosis; alcoholic liver disease; diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; neurofilament-F; presenilin I; presenilin II; cellular tumour antigen; glial fibrillary acidic protein; GRP; p53; semaphorin III; HUPF-1; bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; hMGp-C; NSP-A; high mobility group protein-C; neuroendocrine specific protein A; ss.			

XX Homo sapiens.
 OS
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACADEMIE ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosvelde FG, Van Leeuwen FW;
 DR WPI; 1998-609901/51.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258bp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases,
 CC especially cancer and a wide range of neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's
 CC disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus
 CC type II and many others listed) or susceptibility to these disorders. The
 CC method allows a definitive diagnosis of Alzheimer's disease in living
 CC patients, at an early stage. It is based on the observation that disease
 CC may be caused by mutations in RNA rather than DNA. The invention
 CC describes the use of neuronal system RNA molecules, specifically
 CC proteins including beta-amyloid precursor protein (beta-APP), the
 CC microtubule associated protein tau and Big Tau, ubiquitin B,
 CC apolipoprotein E, microtubule associated protein 2 (MAP2),
 CC neurofilament-L, neurofilament-M, neurofilament-F, prealbumin I,
 CC prealbumin II, glial fibrillary acidic protein (GFAP), the cellular
 CC tumor antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene,
 CC semaphorin III, HUPF-1, high mobility group protein-C (HMGp-C) and
 CC neuroendocrine specific protein A. This sequence encodes the wild type
 CC and mutant protein fragments represented in AAY21264-Y21348.

XX
 SQ Sequence 2530 BP; 786 A; 518 C; 576 G; 650 T; 0 other;

Query Match 92.6%; Score 2508; DB 19; Length 2530;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACGACATGGGCTGGTAACTAGAGATTGTCTCTTTCTGGGAGATTAATTACACG 252
 DB 9 CTGACGACATGGGCTGGTAACTAGAGATTGTCTCTTTCTGGGAGATTAATTACACG 68
 QY 253 AAGACCAACTATCAGAAATGGGAAAGCAATGTGCCAAGCTGMAATTCTTCAAAAGA 312
 DB 69 AAGACCAACTATCAGAAATGGGAAAGCAATGTGCCAAGCTGMAATTCTTCAAAAGA 128
 QY 313 AATGTGGAATCCAAATATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATA 372
 DB 129 AATGTGGAATCCAAATATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATA 188
 QY 373 TACCTTCTTTTGAATGAGAAAGGAGTGTATGTTGAGCAAAAGATCAATATT 432
 DB 189 TACCTTCTTTTGAATGAGAAAGGAGTGTATGTTGAGCAAAAGATCAATATT 248
 QY 433 TTGATTCGACCTGGTTAAATCAAGATTTTCAAAAGATTTGTGTGCCAGTATCTTACG 492
 DB 249 TTGATTCGACCTGGTTAAATCAAGATTTTCAAAAGATTTGTGTGCCAGTATCTTACG 308

QY 493 CAGAGAGATGATCAAGTGGCTGGAAGAACATCTGAAAGATGTGTAATTTCAT 552
 DB 309 CAGAGAGATGATCAAGTGGCTGGAAGAACATCTGAAAGATGTGTAATTTCAT 368
 QY 553 CAGGTAATTAAGGATTAATCAAGTGTGTAAGCTGTGTGGAAGGAGGCTTTTCA 612
 DB 369 CAGGTAATTAAGGATTAATCAAGTGTGTAAGCTGTGTGGAAGGAGGCTTTTCA 428
 QY 613 TCCATTTGACCTTCAATTTGAATTTGACATCATCTGAGGCAATATTTTAACTGGA 672
 DB 429 TCCATTTGACCTTCAATTTGAATTTGACATCATCTGAGGCAATATTTTAACTGGA 488
 QY 673 GAACCTCAATTTTGAAGAACGCGGTGGGAAGTCCATATGACCTTAAGCTGTGACAGC 732
 DB 489 GAACCTCAATTTTGAAGAACGCGGTGGGAAGTCCATATGACCTTAAGCTGTGACAGC 548
 QY 733 ATCCCTTTTAAATAGATGAGAAATTAATCTGGAATCTGGAATCTGAGGCGAGA 792
 DB 549 ATCCCTTTTAAATAGATGAGAAATTAATCTGGAATCTGAGGCGAGAG 608
 QY 793 CTTGCTATCTTCCGAATCTTGGGACACACCATCAAGACAGACAGATGATTC 852
 DB 609 CTTGCTATCTTCCGAATCTTGGGACACACCATCAAGACAGACAGATGATTC 668
 QY 853 CAGTGGCTCAATGATCCAAAGTTCAATTAAGTCCACCTCATCTGAGAGTGAATCC 912
 DB 669 CAGTGGCTCAATGATCCAAAGTTCAATTAAGTCCACCTCATCTGAGAGTGAATCC 728
 QY 913 TGAAGATACAAAGTATCTTTTCTTCCGTAATGCAATAGATGAGAACTCTGG 972
 DB 729 TGAAGATACAAAGTATCTTTTCTTCCGTAATGCAATAGATGAGAACTCTGG 788
 QY 973 AAAGCTCAACGCTAATTAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAATTAAG 1032
 DB 789 AAAGCTCAACGCTAATTAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAATTAAG 848
 QY 1033 TCTGTGTAATTAATGAGCAATCTTCCAAAGCTGTGTAATTTGCTCAGTCCAGGTCC 1092
 DB 849 TCTGTGTAATTAATGAGCAATCTTCCAAAGCTGTGTAATTTGCTCAGTCCAGGTCC 908
 QY 1093 AAATGCAATGACACTCAATTTTATGATGAATGCAATGATGATGATGATGATGATGAT 1152
 DB 909 AAATGCAATGACACTCAATTTTATGATGAATGCAATGATGATGATGATGATGATGAT 968
 QY 1153 TCTTAAATTAATCAGTGTATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1212
 DB 969 TCTTAAATTAATCAGTGTATTAAGTGTGATGATGATGATGATGATGATGATGATGAT 1028
 QY 1213 AGCCGTGTATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
 DB 1029 AGCCGTGTATGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
 QY 1273 CAGGATGAGCCCAATCTATCAATGGGTGCTTATCAAGAAAGATGCTCTATCCAGGCC 1332
 DB 1089 CAGGATGAGCCCAATCTATCAATGGGTGCTTATCAAGAAAGATGCTCTATCCAGGCC 1148
 QY 1333 AGGAATCTTCCCAAGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1392
 DB 1149 AGGAATCTTCCCAAGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1208
 QY 1393 TGTATTAACCTTTTGAAGATGATCAATCCAGCATGATCAATCCAGTGTTCATGAACA 1452
 DB 1209 TGTATTAACCTTTTGAAGATGATCAATCCAGCATGATCAATCCAGTGTTCATGAACA 1268
 QY 1453 TCGCCCAATAGTATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
 DB 1269 TCGCCCAATAGTATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1328
 QY 1513 AGTGAATCAGAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
 DB 1329 AGTGAATCAGAAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388

1573 TCTTAAAGTAGTTTCAATTCCTTAAGAGACCTTGATGATTGTAAGAGGTTCTGCTGGA 1632
 1389 TCTTAAAGTAGTTTCAATTCCTTAAGAGACCTTGATGATTGTAAGAGGTTCTGCTGGA 1448
 1633 AGAATGACAGTTTGGGGAACCGAGCTGATTTTCAAGATGAGGCTTCCACATACGA 1692
 1449 AGAATGACAGTTTGGGGAACCGAGCTGATTTTCAAGATGAGGCTTCCACATACGA 1508
 1693 GCAACACTATATATTTGTTTCAACCGGCTGGGTTCCAGCTCCCTTTACACCGGTGGA 1752
 1509 GCAACACTATATATTTGTTTCAACCGGCTGGGTTCCAGCTCCCTTTACACCGGTGGA 1566
 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTTACTGTGTTGGGA 1812
 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTTACTGTGTTGGGA 1628
 1813 TGTGTTCTGATGTTTCTGCTATTTTCCACTGCAAGAGACGCAACAGCAAGATAT 1872
 1629 TGTGTTCTGATGTTTCTGCTATTTTCCACTGCAAGAGACGCAACAGCAAGATAT 1688
 1873 AAGAAATGAGAGCCCACTGACTCAGCTGTTAGACTTACCATGATATATCAGATGGCA 1932
 1689 AAGAAATGAGAGCCCACTGACTCAGCTGTTAGACTTACCATGATATATCAGATGGCA 1748
 1933 CAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAGACATTTTGGAAATGCGAG 1992
 1749 CAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAGACATTTTGGAAATGCGAG 1808
 1993 TCCGAAGTCCGAGAGAGGCTGTGTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAA 2052
 1809 TCCGAAGTCCGAGAGAGGCTGTGTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAA 1868
 2053 AGAAGAGATCAGAGTGTGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG 2112
 1869 AGAAGAGATCAGAGTGTGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG 1928
 2113 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCGAGTGGGAAACATGGGTTCTATCA 2172
 1929 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCGAGTGGGAAACATGGGTTCTATCA 1988
 2173 AACTCTTTTAAAGTAAACCTCTGGAAGTATTCAGACAGAGATTTGGAAGAACTTTCTCA 2232
 1989 AACTCTTTTAAAGTAAACCTCTGGAAGTATTCAGACAGAGATTTGGAAGAACTTTCTCA 2048
 2233 TAAAGATGATGATGAGATGAGTGTCTTAAGACCAAGAAATGTCATATAGATGACACTTAA 2292
 2049 TAAAGATGATGATGAGATGAGTGTCTTAAGACCAAGAAATGTCATATAGATGACACTTAA 2108
 2293 CCAGAGAGTCTGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGAT 2352
 2109 CCAGAGAGTCTGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGAT 2168
 2353 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAACCTCCGAAAGGCCGAG 2412
 2169 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAACCTCCGAAAGGCCGAG 2228
 2413 ACATACCCCGAGGAAACATTAACAATGAGACCTTAAGAAGAAATTAAGAAAGGTAGAA 2472
 2229 ACATACCCCGAGGAAACATTAACAATGAGACCTTAAGAAGAAATTAAGAAAGGTAGAA 2288
 2473 CAGAGAGACCAAGATTTGAGAGGACCAAGAGTGTCTGAGCTGATTAACCTCTTGA 2532
 2289 CAGAGAGACCAAGATTTGAGAGGACCAAGAGTGTCTGAGCTGATTAACCTCTTGA 2348
 2533 AACCTCAAAACAAGTGAAGAACTTGCTTAACAATTAACCTGAAAAACAATTCATATACAT 2592
 2349 AACCTCAAAACAAGTGAAGAACTTGCTTAACAATTAACCTGAAAAACAATTCATATACAT 2408
 2593 GAACTTTTTCATGATGATGATGATGATTAACAATGAGGAAATTAAGAGCTGATTCGA 2652
 2409 GAACTTTTTCATGATGATGATGATGATTAACAATGAGGAAATTAAGAGCTGATTCGA 2468
 2653 CCAATTAATAATTAATCAATGATGATGATTCCTTAATAGGCTTTT 2700

DB 2469 CCAATTAATAATTAATCAATGATGATGATTCCTTAATAGGCTTTT 2516
 RESULT 3
 ABS76512
 ID ABS76512 standard; cDNA; 2530 BP.
 XX AC ABS76512;
 XX 11-DEC-2002 (first entry)
 XX
 DE cDNA encoding human ovarian cancer marker M473.
 XX
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nontuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200271928-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 14-MAR-2002; 2002W0-US07826.
 XX
 PR 14-MAR-2001; 2001US-276025P.
 XX
 PR 14-MAR-2001; 2001US-276025P.
 PR 10-AUG-2001; 2001US-311732P.
 PR 19-SEP-2001; 2001US-323580P.
 PR 26-SEP-2001; 2001US-324967P.
 PR 26-SEP-2001; 2001US-325102P.
 PR 26-SEP-2001; 2001US-325149P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Monahan JF, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG,
 PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB,
 PI Baet RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX
 DR WPI: 2002-723277/78.
 XX
 DR P-PSDB; ABG95413.
 XX
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient -
 XX
 XX
 XX Disclosure; Page 400; 481pp; English.
 XX
 PS The present invention relates to a new method for assessing whether a
 PS patient is afflicted with ovarian cancer. The method involves comparing
 PS the expression level of a marker in a patient sample and the normal level
 PS of expression of the marker in a control non-ovarian cancer sample, where
 PS the marker is selected from 363 cancer markers described in the
 PS specification. The method of the invention is useful in diagnosing or
 PS characterizing cancer, in detecting the presence of cancer as early as
 PS possible, and the recurrence of ovarian cancer. The method may also be of
 PS particular use with patients having an enhanced risk of developing
 PS ovarian cancer (e.g. patients having a familial history of ovarian
 PS cancer). The cancer markers may be used in the management and treatment
 PS of e.g. brain and central nervous system disorders (e.g. bacterial and
 PS viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 PS disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 PS inflammations (e.g. bacterial or viral meningitis or encephalitis),
 PS testicular disorders (e.g. nontuberculous granulomatous orchitis),
 PS connective tissue disorders, or heart disorders (e.g. ischaemic heart
 PS disease or atherosclerosis). The compositions and methods may also be
 PS used in assessing the histological type of neoplasm associated with

CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present nucleic acid sequence encodes one of the ovarian cancer markers
CC described in the invention.

XX Sequence 2530 BP; 786 A; 518 C; 576 G; 650 T; 0 other;

Query Match 92.6%; Score 2508; DB 24; Length 2530;

Best Local Similarity 100.0%; Pred.No.0;

Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 CTGACGACATGGGCTGGTAACTAGATTGTCTCTTTCTGGGAGTATTAATTACAC 252
9 CTGACGACATGGGCTGGTAACTAGATTGTCTCTTTCTGGGAGTATTAATTACAC 68
253 AAGAGCAAACTATCAGAAATGGGAAAGCAATGTGCCAAGGCTGAAATTAATCTCA 312
69 AAGAGCAAACTATCAGAAATGGGAAAGCAATGTGCCAAGGCTGAAATTAATCTCA 128
313 AATGTGGAAATCCAAATATGTATCACTTTCAATGGCTGGCCCAAGCTCCAGTTATCA 372
129 AATGTGGAAATCCAAATATGTATCACTTTCAATGGCTGGCCCAAGCTCCAGTTATCA 188
372 TACCTTCTTTGGATGAGAAAGGAGTGGCTGTATGTGGAGCAAGATCAATAT 432
189 TACCTTCTTTGGATGAGAAAGGAGTGGCTGTATGTGGAGCAAGATCAATAT 248
433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTAA 492
249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTAA 308
433 CAGAAGAGTGAATGCAATGGGCTGGAAAGACATCCGTAAGAAATGTGCTAATTTCA 552
309 CAGAAGAGTGAATGCAATGGGCTGGAAAGACATCCGTAAGAAATGTGCTAATTTCA 368
553 CAGGACTCTTAAGCATATATCAGACTCATTTGACGCTGTGAAACGGGGCTTTTCA 612
369 CAGGACTCTTAAGCATATATCAGACTCATTTGACGCTGTGAAACGGGGCTTTTCA 428
613 TCCAAATTTGACCTTCAATTTGAATTTGAATCACTCTGAGAGCAATATTTTAAAGCTGA 672
429 TCCAAATTTGACCTTCAATTTGAATTTGAATCACTCTGAGAGCAATATTTTAAAGCTGA 488
673 GAACTCACATTTTGAAGAAAGGCGCTGGGAAAGTCCATATGACCCCTAAGCTGTGACAC 732
489 GAACTCACATTTTGAAGAAAGGCGCTGGGAAAGTCCATATGACCCCTAAGCTGTGACAC 548
733 ATCCCTTTAATAGATGAGAAATTAATCTCTGGAATCTGCACTGATTTATGGGGGAGA 792
549 ATCCCTTTAATAGATGAGAAATTAATCTCTGGAATCTGCACTGATTTATGGGGGAGA 608
793 CTTTGTCTATCTTCCGAATCTTTGGGACCAACCAATCAGAGACAGACAGATGATTC 852
609 CTTTGTCTATCTTCCGAATCTTTGGGACCAACCAATCAGAGACAGACAGATGATTC 668
853 CAGGTGGCTCATGATCCAAAGTTCTATAGTCCCACTCATCTCAGAGAGTGAATATC 912
669 CAGGTGGCTCATGATCCAAAGTTCTATAGTCCCACTCATCTCAGAGAGTGAATATC 728
913 TGAAGATGCAAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGAGAGACATCTGG 972
729 TGAAGATGCAAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGAGAGACATCTGG 788
973 AAAAGCTACTCAGCTAGATAGTTCAGATATGCAAGAAATGCTTTGGAGGGCACAAG 1032
789 AAAAGCTACTCAGCTAGATAGTTCAGATATGCAAGAAATGCTTTGGAGGGCACAAG 848
1033 TCTGTGTAATTAATGCAAACTTCTCAAAGCTGTCTGATTTGTCTGAGTCCAGGTTC 1092
849 TCTGTGTAATTAATGCAAACTTCTCAAAGCTGTCTGATTTGTCTGAGTCCAGGTTC 908

1093 AAATGGCATTTGACACTATTTTGTATGAACTGAGAGTATTTCTTAATGAATTTAAAGA 1152
909 AAATGGCATTTGACACTATTTTGTATGAACTGAGAGTATTTCTTAATGAATTTAAAGA 968
1153 TCTTAAAAATCCAGTTGTATATGAGTGTATTCAGACTTCCAGTAACTTTTCAAGGATTC 1212
969 TCTTAAAAATCCAGTTGTATATGAGTGTATTCAGACTTCCAGTAACTTTTCAAGGATTC 1028
1213 AGCCGATGTATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
1029 AGCCGATGTATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
1273 CAGGATGAGACCCCACTATCAATGGGCTCTTATCAAGAAAGTCCCTATCCAGGCT 1332
1089 CAGGATGAGACCCCACTATCAATGGGCTCTTATCAAGAAAGTCCCTATCCAGGCT 1148
1333 AGGAACTTGTCCAGCAAAACATTTGGTGTGTTTGTACCTCTACAAAGACCTTCTGATGA 1392
1149 AGGAACTTGTCCAGCAAAACATTTGGTGTGTTTGTACCTCTACAAAGACCTTCTGATGA 1208
1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATGATGATGATGATGATGATGATGATGAT 1452
1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATGATGATGATGATGATGATGATGATGAT 1268
1453 TCGCCCAATATGATCAAAACGATATTAATTAATCAATTTTACAAATTTGCTGATGACCG 1512
1269 TCGCCCAATATGATCAAAACGATATTAATTAATCAATTTTACAAATTTGCTGATGACCG 1328
1513 AGTGAATGCAAGATGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
1329 AGTGAATGCAAGATGAGACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
1573 TCTTAAGTATGTTTCAATTTCTTAAGAGAGACCTTGGTATGATTAAGAGAGTTCGCTGGA 1632
1389 TCTTAAGTATGTTTCAATTTCTTAAGAGAGACCTTGGTATGATTAAGAGAGTTCGCTGGA 1448
1633 AGAATGACAGTTTTCGGGAAACCGACTGCTATTTTCAAGCAATGAGCTTTCACATTAACA 1692
1449 AGAATGACAGTTTTCGGGAAACCGACTGCTATTTTCAAGCAATGAGCTTTCACATTAACA 1508
1693 GCAACACTATATATGTTTCAACGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA 1752
1509 GCAACACTATATATGTTTCAACGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA 1568
1753 TATTTAAGGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
1569 TATTTAAGGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
1813 TGTTCCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAACAGACGCAAGATAT 1872
1629 TGTTCCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAACAGACGCAAGATAT 1688
1873 AAGAAATGAGACCCCACTGATCACTGTTTCACTTACACATGATATCACTATGCGCA 1932
1689 AAGAAATGAGACCCCACTGATCACTGTTTCACTTACACATGATATCACTATGCGCA 1748
1933 CAGCCCTGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
1749 CAGCCCTGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
1993 TCCGAAGTTCGAGAGAGGCTGTCTATTTGCAATTCAGAGGCGGAAATGGAAGCGGAA 2052
1809 TCCGAAGTTCGAGAGAGGCTGTCTATTTGCAATTCAGAGGCGGAAATGGAAGCGGAA 1868
2053 AGAAGATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
1869 AGAAGATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
2113 TCTTAAAGAGAGATTCAGGCAATTAATCTGCAATGCGGTGGAACATGGGTTCAATCA 2172
1929 TCTTAAAGAGAGATTCAGGCAATTAATCTGCAATGCGGTGGAACATGGGTTCAATCA 1988

```

Oy 2173 AACTCTTCTTAAGTAACCTCGGAAGTCATTGACACAGACATTGGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTAACCTCGGAAGTCATTGACACAGACATTGGAGAACTTCTTCA 2048
Oy 2223 TAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCATATGACATGACCTTAG 2292
Db 2049 TAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCATATGACATGACCTTAG 2108
Oy 2223 CCAGAGAGCTGCTGACAGAGACTTCATGACGCTCAACCAACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGAGCTGCTGACAGAGACTTCATGACGCTCAACCAACCCCAATCTCAACAGAT 2168
Oy 2353 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCAAAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCAAAAGGCCAGG 2228
Oy 2413 ACATACCCCGAGGAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAAA 2472
Db 2229 ACATACCCCGAGGAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAAA 2288
Oy 2473 CAGAGAGACCAAGATTTGAGAGGGCAACCAAGAGTGTGAGCTGATTAACCTCTAGA 2532
Db 2289 CAGAGAGACCAAGATTTGAGAGGGCAACCAAGAGTGTGAGCTGATTAACCTCTAGA 2348
Oy 2533 AACCTCAAAACAAGTAAGAACTTGCCTAGACAATACTGAAAAACAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAAGTAAGAACTTGCCTAGACAATACTGAAAAACAATGCAATATACAT 2408
Oy 2553 GAACCTTTTTCATGATGATTAATGATGATTTTCAATGATGGAATTAAGCTGACTTCA 2652
Db 2409 GAACCTTTTTCATGATGATTAATGATGATTTTCAATGATGGAATTAAGCTGACTTCA 2468
Oy 2653 CCAATTAATAATTAATCATGATGATTAATGATGATTTTCTTAAATGAGCTTTT 2700
Db 2469 CCAATTAATAATTAATCATGATGATTAATGATGATTTTCTTAAATGAGCTTTT 2516

```

RESULT 4
AA087442
ID AA087442 standard; cDNA, 2601 BP.

AC AA087442;
XX
XX 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
XX Human semaphorin III cDNA.
DE
XX Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KV variola major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection; dr.

OS Homo sapiens.

XX
XX
XX Key Location/Qualifiers
FT CDS 16..2331
FT /tag= a
FT /product= human semaphorin III
FT

XX W09507706-A1.

XX 23-MAR-1995.

XX 13-SEP-1994; 94WO-US10151.

XX 13-SEP-1993; 93US-0121713.

XX (REGC) UNIV CALIFORNIA.

XX Bentley DR, Goodman CS, Kolodkin AL, Matthes D,
PI O'Connor T,
XX

DR WPI; 1995-131177/17.
DR P-PSDB; AAR71380.
XX
XX New class of semaphorin peptide(s) and polypeptide(s) - are
PT potent modulators of nerve cell growth and regeneration
XX
XX
XX Example 2; Page 60-63; 101pp; English.

CC The sequence of the cDNA encoding the human semaphorin III protein.
CC The proteins encoded by the grasshopper semaphorin I (AA087441), human
CC semaphorin III, vaccinia virus semaphorin IV (AA087443), Drosophila
CC semaphorin I and II (AA087444-5), Tribolium semaphorin I (AA087446) or
CC variola major (smallpox) virus semaphorin IV (AA087447) genes were used
CC to generate a series of peptides (AAR70370-R70418), which retain
CC semaphorin receptor binding activity. The semaphorin derived or
CC semaphorin receptor derived peptides are potent modulators of nerve cell
CC growth, immune responsiveness and viral pathogenesis. They can be used
CC in diagnosis and treatment of neurological disease and
CC neuro-regeneration, immune modulation and diagnosis and treatment of
CC viral and oncological infection and diseases.
CC (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 2601 BP; 809 A; 533 C; 593 G; 666 T; 0 other;

Query Match 92.6%; Score 2508; DB 16; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 193 CTGACAGATGGGCTGGTTAACTAGATTTGCTGCTTTCTGGGAGATTAATTACACG 252
Db 9 CTGACAGATGGGCTGGTTAACTAGATTTGCTGCTTTCTGGGAGATTAATTACACG 68
Oy 253 AAGGCAACATCATGATGAGAGAGAAACAATGSCCAAGGCTGAATTAATCTCAAGA 312
Db 69 AAGGCAACATCATGATGAGAGAGAAACAATGSCCAAGGCTGAATTAATCTCAAGA 128
Oy 313 AATGTTGAATCCAAACATGATGATCACTTCAATGCTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTTGAATCCAAACATGATGATCACTTCAATGCTGGCCAAAGCTCCAGTTATCA 188
Oy 373 TACCTTCCTTTGGATGAGAAAGGAGTGTATGTTGAGCAAAAGATCAATATT 432
Db 189 TACCTTCCTTTGGATGAGAAAGGAGTGTATGTTGAGCAAAAGATCAATATT 248
Oy 433 TTCAATTCGACCTGTTAATCAAGATTTTCAAAAGATGTTGGCCAGATTTTACAC 492
Db 249 TTCAATTCGACCTGTTAATCAAGATTTTCAAAAGATGTTGGCCAGATTTTACAC 308
Oy 493 CAGAAGAGATGAAGCAAGTGGCTGGAAAAAGACATCCTGAAAGATGCTAATTTTCA 552
Db 309 CAGAAGAGATGAAGCAAGTGGCTGGAAAAAGACATCCTGAAAGATGCTAATTTTCA 368
Oy 553 CAAGTACTTAAGCATTAATCACTCACTCTGTGAGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGTACTTAAGCATTAATCACTCACTCTGTGAGCTGTGGAACGGGGCTTTTCA 428
Oy 613 TTCAATTTGACCCATCAATTAATGAGATCAATCCTGAGGAGCAATATTTTAAGCTGGA 672
Db 429 TTCAATTTGACCCATCAATTAATGAGATCAATCCTGAGGAGCAATATTTTAAGCTGGA 488
Oy 673 GAACCTCAATTTTGAAGAAAGCGCCGTGGGAAGAGTCATATGACCTTAAGCTGTGACAGC 732
Db 489 GAACCTCAATTTTGAAGAAAGCGCCGTGGGAAGAGTCATATGACCTTAAGCTGTGACAGC 548
Oy 733 ATCCCTTTTAATGATGAGATTAATCACTGTAAGCTGACGCTGATTTTATGGGCGAGA 792
Db 549 ATCCCTTTTAATGATGAGATTAATCACTGTAAGCTGACGCTGATTTTATGGGCGAGA 608
Oy 793 CTTTGCTATCTTCGGAATCTTTGGGACCAACACCAATCAGAGACAGAGCATGATTC 852
Db 609 CTTTGCTATCTTCGGAATCTTTGGGACCAACACCAATCAGAGACAGAGCATGATTC 668
Oy 853 CAGGTGCTCAATGATCAAAAGTTCAATTAGTGCACCACTCATCTCAGAGAGTGAATCC 912

```


PF 15-OCT-1993; 93US-0136922.
XX
PR 15-OCT-1993; 93US-0136922.
XX
PA (UTPE-) UNIV PENNSYLVANIA.
P1 Luo Y, Raper JA;
XX MPI, 1995-193478/25.
DR P-PSDB; AAR74175.
XX
PT New antibody to human collapsin - used to inhibit the activity of
PT collapsin, to induce neurite out-growth and to treat individuals with
PT nerve damage.
PS Disclosure: Columns 11-16; 11pp; English.
CC Human collapsin and its encoding nucleic acid may be used to
CC identify agents which modulate the ability of human collapsin to
CC collapse the growth cone of neurons. An antibody capable of
CC specifically binding at least a portion of the collapsin protein can
CC be used to purify human collapsin and to inhibit the activity of the
CC protein. It can be used to induce neurite outgrowth by neuronal
CC cells and to treat individuals suffering from nerve damage.
SQ Sequence 1481 BP; 454 A; 299 C; 345 G; 383 T; 0 other;
Query Match 46.3%; Score 1253; DB 16; Length 1481;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 904 TGACATCTCTGAAGATGACAAAGTACTTTCTTCCTGGAATGCAATGATGAGA 963
DB 58 TGACATCTCTGAAGATGACAAAGTACTTTCTTCCTGGAATGCAATGATGAGA 117
OY 964 ACACTCTGGAAGGCTACTCAAGCTGAGATGATGCAATGCAATGATGAGA 1023
DB 118 ACACTCTGGAAGGCTACTCAAGCTGAGATGATGCAATGCAATGATGAGA 177
OY 1024 GCACGAGTCTGAGATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTGCT 1083
DB 178 GCACGAGTCTGAGATTAATGAGCAACATTCCTCAAGCTGCTGATTTGCTGCT 237
OY 1084 GCCAGGTCCAATGAGCATTCACATCTTTGATGATGATGATGATGATGATGAT 1143
DB 238 GCCAGGTCCAATGAGCATTCACATCTTTGATGATGATGATGATGATGATGAT 297
OY 1144 CTTTAAAGATCTTAAATTCAGTTGATGATGATGATGATGATGATGATGATGAT 1203
DB 298 CTTTAAAGATCTTAAATTCAGTTGATGATGATGATGATGATGATGATGATGAT 357
OY 1204 CAAGGATCAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
DB 358 CAAGGATCAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
OY 1264 ATATGCCCAAGGATGAGCCCACTATCAATGGGTGCTTATCAAGGAAGTCCCTA 1323
DB 418 ATATGCCCAAGGATGAGCCCACTATCAATGGGTGCTTATCAAGGAAGTCCCTA 477
OY 1324 TCCACGGCCAGGAATCTTCCAGCAAAATTTGGTGGTTTGACTTCAACAAAGACT 1383
DB 478 TCCACGGCCAGGAATCTTCCAGCAAAATTTGGTGGTTTGACTTCAACAAAGACT 537
OY 1384 TCCATGATGATGATTAACCTTTGCAAGAGTCAATCAACCTGATCAATGATGAT 1443
DB 538 TCCATGATGATGATTAACCTTTGCAAGAGTCAATCAACCTGATCAATGATGAT 597
OY 1444 TATGAACATCGCCCAATAGTATCAAAACGATGATTAATTAATTAATTAATTAAT 1503
DB 598 TATGAACATCGCCCAATAGTATCAAAACGATGATTAATTAATTAATTAATTAAT 657
OY 1504 CGTAGACCGAGTGTGATGAGAAAGATGACAGTATGATGATGATGATGATGATGAT 1563

DB 658 GTTAGACCGAGTGTGATGAGAAAGATGACAGTATGATGATGATGATGATGATGAT 717
OY 1564 TGAGACCGGTTCTTAAAGTATGATTAATCTTAAAGAGACTGATGATGATTAAGAGAGT 1623
DB 718 TGAGACCGGTTCTTAAAGTATGATTAATCTTAAAGAGACTGATGATGATTAAGAGAGT 777
OY 1624 TCTGCTGGAAGAAATGACAGTTTTCGGAACCCGACTGCTATTTTCAAGCAATGAGCTTTC 1683
DB 778 TCTGCTGGAAGAAATGACAGTTTTCGGAACCCGACTGCTATTTTCAAGCAATGAGCTTTC 837
OY 1684 CACTTAACGACGACAACTATATATGTTGTTCAAGCGCTGGGGTTCCTCCAGCTTCTTACA 1743
DB 838 CACTTAACGACGACAACTATATATGTTGTTCAAGCGCTGGGGTTCCTCCAGCTTCTTACA 897
OY 1744 CCGGTGTGATTAATTAACGGGAAAGGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCT 1803
DB 898 CCGGTGTGATTAATTAACGGGAAAGGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCT 957
OY 1804 TGCTTGGGATGTTCTGCAATGTTCTGCTATTTTCCACTGCAAAAGAGAGCAGCAGAG 1863
DB 958 TGCTTGGGATGTTCTGCAATGTTCTGCTATTTTCCACTGCAAAAGAGAGCAGCAGAG 1917
OY 1864 ACAAGATTAAGAAATGAGAGCCCACTGACTCTGCTTCACTTCACTTCACTTCACTTCA 1923
DB 1918 ACAAGATTAAGAAATGAGAGCCCACTGACTCTGCTTCACTTCACTTCACTTCACTTCA 1977
OY 1924 CCATGGCCACAGCCCTGAAGAGAGATCATCTATGATGATGATGATGATGATGATGATGAT 1983
DB 1078 CCATGGCCACAGCCCTGAAGAGAGATCATCTATGATGATGATGATGATGATGATGATGAT 1137
OY 1984 GGAATGAGTCCGAGTCCGAGAGAGAGGCTGCTGCTATTTGCAATTCAGAGCCGAAATGA 2043
DB 1138 GGAATGAGTCCGAGTCCGAGAGAGAGGCTGCTGCTATTTGCAATTCAGAGCCGAAATGA 1197
OY 2044 AGACGCAAAAGAGAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103
DB 1198 AGACGCAAAAGAGAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1257
OY 2104 GCTAGTATGCTTACACAGAAAGATTCAGGCAATTAATTCAGGCTGAGGAGAGAGAG 2163
DB 1258 GCTAGTATGCTTACACAGAAAGATTCAGGCAATTAATTCAGGCTGAGGAGAGAGAG 1317
OY 2164 GTTCATACAAATCTTCTTAAAGTAACTCTGAGGATGATGATGATGATGATGATGAT 2207
DB 1318 GTTCATACAAATCTTCTTAAAGTAACTCTGAGGATGATGATGATGATGATGATGAT 1361
RESULT 6
ABS45213
ID ABS45213 standard; DNA; 456 BP.
XX
AC ABS45213;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 20203.
XX
KW Human, single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
EN WO200157273-A2.
XX
PF 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -
 PS Claim 4; SEQ ID No 20203; 658bp; English.
 CC The invention relates to a single exon nucleic acid probe (SENP) (i) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (i) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WFO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 456 BP; 161 A; 98 C; 108 G; 89 T; 0 other;
 Query Match 15.0%; Score 405; DB 23; Length 456;
 Best Local Similarity 99.8%; Prod 2.6e-189;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2060 ATCAGAGTGATGATCATATCATCAGACACATCAAGCCCTTCTGCTACGTAGCTTACAA 2119
 DB 1 ATCAGAGTGATGATCATATCATCAGACACATCAAGCCCTTCTGCTACGTAGCTTACAA 60
 QY 2120 CAGAGGATTCAGGCAATACCTCTGCACTGCGGTGGAACATGGGTCTACATAAACTCT 2179
 DB 61 CAGAGGATTCAGGCAATACCTCTGCACTGCGGTGGAACATGGGTCTACATAAACTCT 120
 QY 2180 CTTAAGGTAACTCTGGAAGTCAATGACACAGAGCATTTTGAAGAACTTCTTCAATAAGAT 2239
 DB 121 CTTAAGGTAACTCTGGAAGTCAATGACACAGAGCATTTTGAAGAACTTCTTCAATAAGAT 180
 QY 2240 GATGATGAGATGGCTCTAAGACCAAAAGATGTCCAATAGCATGACACTTACCCAGAG 2299
 DB 181 GATGATGAGATGGCTCTAAGACCAAAAGATGTCCAATAGCATGACACTTACCCAGAG 240
 QY 2300 GTCCTGTACAGAGCTTCTATGACGCTCATCAACCCCAATTTCAACGATGATGAG 2359
 DB 241 GTCCTGTACAGAGCTTCTATGACGCTCATCAACCCCAATTTCAACGATGATGAG 300
 QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCAAAACAAGTCTGCAAGGACGAGACATACC 2419
 DB 301 TTCTGTGAACAAGTTTGAAGAAAGGACCAAAACAAGTCTGCAAGGACGAGACATACC 360
 QY 2420 CCAGGAGACAGTACCAATGAGAACCTTACCAAGAAATTAAGAAAGTGAAGAGAGG 2479
 DB 361 CCAGGAGACAGTACCAATGAGAACCTTACCAAGAAATTAAGAAAGTGAAGAGAGG 420
 QY 2480 ACCCAAGATTTGAGAGGCGCCAGGAGTGTCTGA 2515
 DB 421 ACCCAAGATTTGAGAGGCGCCAGGAGTGTCTGA 456
 RESULT 7
 ABS19795
 ID ABS19795 standard; DNA; 456 BP.
 XX

AC ABS19795;
 DR 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe ORF from lung SEQ ID No 19786.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Heremakly-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PP 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 XX
 PS Claim 4; SEQ ID No 19786; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

RESULT 9
ABZ33368
ID ABZ33368 standard; cDNA; 354 BP.
XX
AC ABZ33368;
XX
DT 30-JAN-2003 (first entry)
XX
DE Human colon tumour cDNA for clone R0093; F06 SEQ ID NO: 736.
XX
KM Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
KW tumour; immune response; immunostimulant; cytostatic; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283070-A2.
XX
PD 24-OCT-2002.
XX
PF 09-APR-2002; 2002WO-US11475.
XX
PR 10-APR-2001; 2001US-0833263.
XX
PR 03-AUG-2001; 2001US-0922217.
XX
PR 19-DEC-2001; 2001US-0025380.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD;
PI Skeiky YAW, Fanger GR, Vedvick JS, Carter D;
XX
DR WPI: 2003-067548/06.
XX
PT New polynucleotide, useful for the preparation of a composition for
PT stimulating an immune response against, or treating, cancer -
XX
PS Disclosure; Page 343; 537pp; English.
XX
CC The present invention describes compounds (1) for the immunotherapy and
CC diagnosis of colon cancer. Also described: (1) a method for detecting
CC the presence of cancer in a patient; (2) a method for stimulating and/or
CC expanding T cells specific for a tumour protein; (3) an isolated T cell
CC population comprising T cells prepared by the method of (2); (4) a method
CC for stimulating an immune response in a patient; (5) a method for
CC treating cancer in a patient; and (6) a method for inhibiting the
CC development of cancer in a patient. (1) have immunostimulant and
CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 354 BP; 120 A; 82 C; 79 G; 72 T; 1 other;
XX
Query Match 11.2%; Score 303; DB 25; Length 354;
Best Local Similarity 99.7%; Pred. No. 6,2e-139;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2098 CCTTGTGCTAGCTCTTACACAGAGGATTGAGGCAATTACCTGCGCATGCGGATGA 2157
DB 1 CTTTGTGCTAGCTCTTACACAGAGGATTGAGGCAATTACCTGCGCATGCGGATGA 60
QY 2158 ACATGGTTTCATACAACTCTTAAAGTAAACCTGGAAGTCAATGACAGAGATT 2217
DB 61 ACATGGTTTCATACAACTCTTAAAGTAAACCTGGAAGTCAATGACAGAGATT 120
QY 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCGA 2277
DB 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCGA 180
QY 2278 TAGCATGACACTTACCGAGAGGCTGTGTAAGAGACTTCAAGAGCTCATTAACACACC 2337
DB 181 TAGCATGACACTTACCGAGAGGCTGTGTAAGAGACTTCAAGAGCTCATTAACACACC 240

QY 2338 CAATCTCAACAGATGATGAGTTCTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 2397
DB 241 CAATCTCAACAGATGATGAGTTCTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 300
QY 2398 TCGGCMAAGGCCAGACATACCCCGAGGACAGTACAAATGAGACACTTACA 2451
DB 301 TCGGCMAAGGCCAGACATACCCCGAGGACAGTACAAATGAGACACTTACA 354
XX
RESULT 10
ABZ32213
ID ABZ32213 standard; DNA; 496 BP.
XX
AC ABZ32213;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 7203.
XX
KM Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinemia; hyperlipidaemia; hypercholesterolemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-488898/53.
XX
DR WPI: 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
PS Claim 1; SEQ ID NO 7203; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolemia which
CC is associated with coronary heart disease. ABZ5011-ABZ51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 496 BP; 164 A; 101 C; 92 G; 139 T; 0 other;
XX
Query Match 10.8%; Score 292; DB 23; Length 496;
Best Local Similarity 100.0%; Pred. No. 1,7e-133;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2058 AGATCAGATGATGATCATATCATCAAGACAGATCAAGGCTTCTGCTAGTACTTAC 2117

```

Db 160 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGCCCTTGTGCTACCTAGCTTAC
Oy 2118 AACAGAGGATTCAGGCAATTACCTCTGCGATGCGGTGGAGACATGGGTTCATACAAATC
Db 220 AACAGAGGATTCAGGCAATTACCTCTGCGATGCGGTGGAGACATGGGTTCATACAAATC
Oy 2178 TTCTTAAGTAACTTCGAAAGTATGACACAGACATTTGGAGAACTTCTTCAATAAG
Db 280 TTCTTAAGTAACTTCGAAAGTATGACACAGACATTTGGAGAACTTCTTCAATAAG
Oy 2228 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA
Db 340 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA
Oy 2228 AGGTCTGTACAGAGACTTCAATGACGCTCATGACACCCCAATCTCAACAC
Db 400 AGGTCTGTACAGAGACTTCAATGACGCTCATGACACCCCAATCTCAACAC 451

```

RESULT 11

ABS07290
ID ABS07290 standard; DNA; 496 BP.

AC ABS07290;
XX
DT 19-AUG-2002 (first entry)
XX

DE Human genome-derived single exon probe from lung SEQ ID No 7281.

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW Chronic obstructive pulmonary disease; interstitial lung disease;
KW Familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW Tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW Pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
KW Pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW Primary ciliary dyskinesia; pulmonary hypertension;
KW Hyaline membrane disease.

OS Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples -

XX Claim 1; SEQ ID No 7281; 634bp; English.

CC probes; the novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 496 BP; 164 A; 101 C; 92 G; 139 T; 0 other;

XX Query Match 10.8%; Score 292; DB 24; Length 496;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-143; Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 2058 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGCCCTTGTGCTAGTACTAC 2117
Db 160 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGCCCTTGTGCTAGTACTAC 219
Oy 2118 AACAGAGGATTCAGGCAATTACCTCTGCGATGCGGTGGAGACATGGGTTCATACAAATC 2177
Db 220 AACAGAGGATTCAGGCAATTACCTCTGCGATGCGGTGGAGACATGGGTTCATACAAATC 279
Oy 2178 TTCTTAAGTAACTTCGAAAGTATGACACAGACATTTGGAGAACTTCTTCAATAAG 2237
Db 280 TTCTTAAGTAACTTCGAAAGTATGACACAGACATTTGGAGAACTTCTTCAATAAG 339
Oy 2228 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA 2297
Db 340 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA 399
Oy 2228 AGGTCTGTACAGAGACTTCAATGACGCTCATGACACCCCAATCTCAACAC 2349
Db 400 AGGTCTGTACAGAGACTTCAATGACGCTCATGACACCCCAATCTCAACAC 451

```

RESULT 12

ABA71205/c
ID ABA71205 standard; DNA; 172 BP.

XX ABA71205;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe. #19510.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

```
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver -
XX PS Claim 4; SEQ ID NO 19510; 639pp + sequence listing; English.
XX CC The invention relates to a single exon nucleic acid probe for
XX CC measuring human gene expression in a sample derived from human fetal
XX CC liver. The single exon nucleic acid probes may be used for predicting,
XX CC measuring and displaying gene expression in samples derived from human
XX CC fetal liver. The present sequence is a single exon nucleic acid
XX CC probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;
XX
XX Query Match 5.9%; Score 160; DB 22; Length 172;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-68;
XX Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 310 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 369
XX DB 160 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 101
XX
XX QY 370 TCATACCTTCCCTTTTGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 429
XX DB 100 TCATACCTTCCCTTTTGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 41
XX
XX QY 430 ATTTTCATTGACCTCGTTAATATCAAGGATTTTCAAAAG 469
XX DB 40 ATTTTCATTGACCTCGTTAATATCAAGGATTTTCAAAAG 1
XX
XX RESULT 13
XX AAK19504/C
XX ID AAK19504 standard; DNA; 172 BP.
XX
XX AC AAK19504;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe SEQ ID NO: 19495.
XX
XX KW Human; brain expressed exon; gene expression analysis; probe;
XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KW epilepsy; cancer; ss.
XX
XX OS Homo sapiens.
XX
```

```
PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00667.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483446/52.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains -
XX PS Example 4; SEQ ID NO: 19495; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention.
XX SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;
XX
XX Query Match 5.9%; Score 160; DB 22; Length 172;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-68;
XX Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 310 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 369
XX DB 160 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 101
XX
XX QY 370 TCATACCTTCCCTTTTGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 429
XX DB 100 TCATACCTTCCCTTTTGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 41
XX
XX QY 430 ATTTTCATTGACCTCGTTAATATCAAGGATTTTCAAAAG 469
XX DB 40 ATTTTCATTGACCTCGTTAATATCAAGGATTTTCAAAAG 1
XX
XX RESULT 14
XX AAK45495/C
XX ID AAK45495 standard; DNA; 172 BP.
XX
XX AC AAK45495;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 20052.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00668.
XX
```

04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR WPI; 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 20052; 658bp + Sequence Listing; English.
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 310 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 101
 OY 370 TCATACCTCTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAAAGATCACAT 429
 DB 100 TCATACCTCTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAAAGATCACAT 41
 OY 430 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 1

RESULT 15

AA151439/C
 ID AA151439 standard; DNA, 172 BP.

AC AA151439;
 XX 17-OCT-2001 (first entry)
 DT
 XX

DE Probe #20125 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 KW
 XX

OS Homo sapiens.
 XX
 XX

PN WO200157272-A2.
 XX
 XX

PD 09-AUG-2001.
 XX
 XX

PE 30-JAN-2001; 2001WO-US00663.
 XX
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR WPI; 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 20125; 654bp; English.
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 310 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 101
 OY 370 TCATACCTCTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAAAGATCACAT 429
 DB 100 TCATACCTCTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAAAGATCACAT 41
 OY 430 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 1

RESULT 16

ABS45181/C
 ID ABS45181 standard; DNA, 172 BP.

AC ABS45181;
 XX 25-FEB-2003 (first entry)
 DT
 XX

DE Human liver single exon probe, SEQ ID NO 20171.

XX Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 XX

OS Homo sapiens.
 XX
 XX

PN WO200157273-A2.
 XX
 XX

PD 09-AUG-2001.
 XX
 XX

PE 30-JAN-2001; 2001WO-US00664.
 XX
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX

DR WPI; 2001-488900/53.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -
 PS Claim 4; SEQ ID No 20171; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

QY Query Match 5.9%; Score 160; DB 23; Length 172;
 DB Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCCAACAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCCAACAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAGATCAAT 429
 DB 100 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAGATCAAT 41

QY 430 ATTTTCATTCGACCTGTTATATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTTATATATCAAGATTTTCAAAAG 1

RESULT 17
 ABS19763/c
 ID ABS19763 standard; DNA; 172 BP.

AC ABS19763;
 XX
 XX
 DT 19-AUG-2002 (first entry)
 XX
 XX
 DE Human genome-derived single exon ORF from lung SEQ ID No 19754.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.
 XX
 XX
 PN WO200186003-A2.
 XX
 XX
 PD 15-NOV-2001.
 XX
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 CC Claim 4; SEQ ID No 19754; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhage, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 XX
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

QY Query Match 5.9%; Score 160; DB 24; Length 172;
 DB Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCCAACAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCCAACAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAGATCAAT 429
 DB 100 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGGAGCAAGATCAAT 41

QY 430 ATTTTCATTCGACCTGTTATATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTTATATATCAAGATTTTCAAAAG 1

```
RESULT 18
ID ABA58707/C
XX ABA58707 standard; DNA; 484 BP.
AC
XX ABA58707;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #7012.
XX
XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 7012; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
SQ
Query Match 5.9%; Score 160; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCGCTTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCGCTTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTCTTTGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 429
DB 398 TCATACCTTCTCTTTGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 339
QY 430 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 299
RESULT 19
ID AAK06842/C
XX AAK06842 standard; DNA; 484 BP.
XX
XX AAK06842;
```

```
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 6833.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 6833; 650pp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancer. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
SQ
Query Match 5.9%; Score 160; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCGCTTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCGCTTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTCTTTGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 429
DB 398 TCATACCTTCTCTTTGGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAT 339
QY 430 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGCACTGGTTAATATCAAGATTTTCAAAAG 299
RESULT 20
ID AAK32561/C
XX AAK32561 standard; DNA; 484 BP.
XX
XX AAK32561;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 7118.
XX
```

KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 OS Homo sapiens.
 XX WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 7118; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 22; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 310 AGAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTA 369
 Db 458 AGAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTA 399
 QY 370 TCATACCTTCCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 429
 Db 398 TCATACCTTCCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 339
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 469
 Db 338 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 299
 XX
 RESULT 21
 AA138386/c
 ID AA138386 standard; DNA; 484 BP.
 XX
 AC AA138386;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #7072 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157272-A2.
 XX

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-48897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 7072; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENPs).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 22; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 310 AGAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTA 369
 Db 458 AGAAATGTTGGAATCCAAATGATGATCACTTTCATGCTGGCCAAAGCTCCAGTTA 399
 QY 370 TCATACCTTCCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 429
 Db 398 TCATACCTTCCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 339
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 469
 Db 338 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 299
 XX
 RESULT 22
 ABS32272/c
 ID ABS32272 standard; DNA; 484 BP.
 XX
 AC ABS32272;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID NO 7262.
 XX
 KW Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX

Db 337 ACATTTTGGATGAGTCCGAAAGTCGAGAGAGCGCTGCTATTGGCAATTCAGAG 396
 Oy 2036 CGAATGAGAGCGGAAAAGAGAG 2059
 Db 397 CGAATGAGAGCGGAAAAGAGAG 420
 RESULT 26
 ABS08426 ID ABS08426 standard; DNA; 446 BP.
 AC ABS08426;
 XX 19-AUG-2002 (first entry)
 XX Human genome-derived single exon probe from lung SEQ ID No 8417.
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS WO200186003-A2.
 PN 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.
 PE 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 PT Claim 1; SEQ ID No 8417; 634bp; English.
 PS The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 446 BP; 142 A; 88 C; 85 G; 131 T; 0 other;
 SQ
 Query Match 5.3%; Score 144; DB 24; Length 446;
 Best Local Similarity 100.0%; Pred. No. 2e-60;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1916 GATATCACCATGCGCACAGCCCTGAAAGAGGAAATCATCTATGCTAGAGAAATAGTACC 1975
 Db 277 GATATCACCATGCGCACAGCCCTGAAAGAGGAAATCATCTATGCTAGAGAAATAGTACC 336
 Oy 1976 ACATTTTGAATGCAATCCGCAAGTCCGACAGAGCGCTGCTATTGGCAATTCAGAGG 2035
 Db 337 ACATTTTGAATGCAATCCGCAAGTCCGACAGAGCGCTGCTATTGGCAATTCAGAGG 396
 Oy 2036 CGAATGAGAGCGGAAAAGAGAG 2059
 Db 397 CGAATGAGAGCGGAAAAGAGAG 420
 RESULT 27
 ID AAK20386
 XX AAK20386 standard; DNA; 123 BP.
 AC AAK20386;
 XX 05-NOV-2001 (first entry)
 DT Human brain expressed single exon probe SEQ ID NO: 20377.
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX Homo sapiens.
 OS WO200157275-A2.
 FN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00667.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.

```
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 20377; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;
XX
SQ
Query Match 4.5%; Score 123; DB 22; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.9e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1937 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCACTCCG 1996
DB 1 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCACTCCG 60
QY 1997 AAGTGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAGGCAAGAAAGAA 2056
DB 61 AAGTGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAGGCAAGAAAGAA 120
QY 2057 GAG 2059
DB 121 GAG 123
RESULT 28
ABS46246
ID ABS46246 standard; DNA; 123 BP.
XX
XX ABS46246;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID No 21236.
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00664.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
```

```
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
XX Claim 4; SEQ ID No 21236; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult
XX liver. (1) may be used for predicting, measuring and displaying gene
XX expression in samples derived from human adult liver. The genes
XX identified may be involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
XX is associated with coronary heart disease. ABS25011-ABS51005 represent
XX human liver single exon nucleic acid probes of the invention.
XX
XX Note: The sequence information for this patent does not appear in the
XX printed specification but was obtained in electronic format directly
XX from WIPO at http://pub.int/pub/published\_pct\_sequences.
XX
SQ
Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;
XX
XX
Query Match 4.5%; Score 123; DB 23; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.9e-50;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1937 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCACTCCG 1996
DB 1 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCACTCCG 60
QY 1997 AAGTGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAGGCAAGAAAGAA 2056
DB 61 AAGTGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAGGCAAGAAAGAA 120
QY 2057 GAG 2059
DB 121 GAG 123
RESULT 29
ABS20840
ID ABS20840 standard; DNA; 123 BP.
XX
XX ABS20840;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon ORF from lung SEQ ID No 20831.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX
```


PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2002-114183/15.
 DR

PT Spatially-addressable set of single exon nucleic acid probes, used to
 measure gene expression in human lung samples -

XX Claim 4; SEQ ID NO 20831; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;

Query Match 4.5%; Score 123; DB 24; Length 123;
 Best Local Similarity 100.0%; Pred. No. 4.9e-50;

Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACAGATTGGAATGCAATGCTCG 1996
 DB 1 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACAGATTGGAATGCAATGCTCG 60

QY 1997 AAGTTCGAGAGAGGCTGCTTATTTGCAATTTCCAGAGCCGAATGAAGCGAAGAA 2056
 DB 61 AAGTTCGAGAGAGGCTGCTTATTTGCAATTTCCAGAGCCGAATGAAGCGAAGAA 120

QY 2057 GAG 2059
 DB 121 GAG 123

RESULT 30
 ID ABA58557/C
 XX ABA58557 standard; DNA; 456 BP.
 AC
 XX ABA58557;
 DT
 XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #6862.

KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -

XX Claim 1; SEQ ID NO 6862; 639bp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 22; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;

Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGAATGCTGCTATTTGCAAGTACTTAAGCATATATATGACTGCTGTCGC 591
 DB 407 GAAGAATGCTGCTATTTGCAAGTACTTAAGCATATATATGACTGCTGTCGC 348

QY 592 CTGTGGAACGGGGGCTTTTCAATTCACCTTACATTTGAATTTGACATCTGGA 651
 DB 347 CTGTGGAACGGGGGCTTTTCAATTCACCTTACATTTGAATTTGACATCTGGA 288

QY 652 GG 653
 DB 287 GG 286

RESULT 31
 ID ABA27590/C
 XX ABA27590 standard; DNA; 456 BP.
 AC ABA27590;

DT	23-JAN-2002	(first entry)
XX		
XX	Probe #6056	for gene expression analysis in human heart cell sample.
DE		
XX		
XX	Human; gene expression; heart; microarray; vascular system; probe;	
KW	cardiovascular disease; hypertension; cardiac arrhythmia;	
KW	congenital heart disease; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157274-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001WO-US00666.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000GB-0024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
P1	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI, 2001-488899/53.	
XX		
FT	Single exon nucleic acid probes for analyzing gene expression in human	
XX	hearts -	
PS	Claim 1; SEQ ID No 6056; 530bp; English.	
XX		
CC	The present invention relates to single exon nucleic acid probes for	
CC	measuring human gene expression in a sample derived from human heart. The	
CC	present sequence is one such probe. The probes may be used for	
CC	predicting, measuring and displaying gene expression in samples derived	
CC	from the human heart via microarrays. By measuring gene expression, the	
CC	probes are useful for predicting, diagnosing, grading, staging,	
CC	monitoring and prognosing diseases of the human heart and vascular system	
CC	e.g. cardiovascular disease, hypertension, cardiac arrhythmias and	
CC	congenital heart disease.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;	
Q0		
	Query Match	4.5%; Score 122; DB 22; Length 456;
	Best Local Similarity	100.0%; Pred. No. 1,5e-49;
	Matches 122; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	532 GAAAGATGTCCTAATTCATCAAGGACTTAAGGCATTAATCAGACTCTTGTAACG	591
DB	407 GAAAGAAATGTCCTAATTCATCAAGGACTTAAGGCATTAATCAGACTCTTGTAACG	348
QY	592 CTGTGGAAACGGGGGCTTTTCATCCAAATTGGACACTTGAATTTGGACATATCTCTGA	651
DB	347 CTGTGGAAACGGGGGCTTTTCATCCAAATTGGACACTTGAATTTGGACATATCTCTGA	288
QY	652 GG 653	
DB	287 GG 286	

RESULT 32

AAK06676/c

ID . AAK06676 standard; DNA; 456 BP.

XX

AAK06676;

IC

[illegible]

KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX Homo sapiens.
XX WO200157276-A2.
XX
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 6929; 658bp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
SQ
Query Match 4.5%; Score 122; DB 22; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.5e-49;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 532 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGTCACTGTAAGC 591
DB 407 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGTCACTGTAAGC 348
QY 592 CTGTGGAACGGGGGCTTTTCATCAATTGCACTTGAATTTGACATCATCTCTGA 651
DB 347 CTGTGGAACGGGGGCTTTTCATCAATTGCACTTGAATTTGACATCATCTCTGA 288
QY 652 GG 653
DB 287 GG 286
RESULT 34
AA138218/c
ID AA138218 standard; DNA; 456 BP.
XX
XX AA138218;
XX
XX 17-OCT-2001 (first entry)
XX
XX DE Probe #6904 used to measure gene expression in human placenta sample.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX OS Homo sapiens.
XX
XX WO200157272-A2.
XX

EP 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 6904; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
SQ
Query Match 4.5%; Score 122; DB 22; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.5e-49;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 532 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGTCACTGTAAGC 591
DB 407 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGTCACTGTAAGC 348
QY 592 CTGTGGAACGGGGGCTTTTCATCAATTGCACTTGAATTTGACATCATCTCTGA 651
DB 347 CTGTGGAACGGGGGCTTTTCATCAATTGCACTTGAATTTGACATCATCTCTGA 288
QY 652 GG 653
DB 287 GG 286
RESULT 35
ABS32080/c
ID ABS32080 standard; DNA; 456 BP.
XX
XX ABS32080;
XX
XX 25-FEB-2003 (first entry)
XX
XX DE Human liver single exon probe, SEQ ID No 7070.
XX
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX OS Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00664.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PA WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -

PS Claim 1; SEQ ID No 7070; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (II) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 23; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGAATGCTAATTTTCATCAGTACTTAAGCATATTAATCAGACTCTTGATGCC 591

DB 407 GAAAGATGCTAATTTTCATCAGTACTTAAGCATATTAATCAGACTCTTGATGCC 348

QY 552 CTGTGACGCGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTTGA 651

DB 347 CTGTGACGCGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTTGA 288

QY 652 GG 653

DB 287 GG 286

RESULT 36
 ABS07155/C
 ID ABS07155 standard; DNA; 456 BP.

AC ABS07155;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID No 7146.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.

PF 04-FEB-2000; 2000US-180112P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PA WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

PS Claim 1; SEQ ID No 7146; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hereditary-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 24; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGAATGCTAATTTTCATCAGTACTTAAGCATATTAATCAGACTCTTGATGCC 591

|||||

Db 407 GAAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 348
 QY 592 CTGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 651
 Db 347 CTGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 288
 QY 652 GG 653
 Db 287 GG 286

RESULT 37

ABA71291/c
 ID ABA71291 standard; DNA; 101 BP.

AC ABA71291;
 XX

DT 01-FEB-2002 (first entry)
 XX

DE Human foetal liver single exon nucleic acid probe #19596.
 XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX

OS Homo sapiens.
 XX

PN WO200157277-A2.
 XX

PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 XX

PR 26-MAY-2000; 2000US-0207456.
 XX

PR 30-JUN-2000; 2000US-0608408.
 XX

PR 03-AUG-2000; 2000US-0632366.
 XX

PR 21-SEP-2000; 2000US-0234687.
 XX

PR 27-SEP-2000; 2000US-0236359.
 XX

PR 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX

PT Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human fetal liver -

PS Claim 4; SEQ ID NO 19596; 639pp + sequence listing; English.
 XX

CC The invention relates to a single exon nucleic acid probe for
 XX measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,
 XX measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid
 XX probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
 XX

Query Match 3.7%; Score 101; DB 22; Length 101;
 XX Best Local Similarity 100.0%; Pred. No. 3.6e-39;

Db 533 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 592
 XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 593 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 592
 XX 101 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 42
 QY 593 TGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 633
 Db 41 TGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 1

RESULT 38
 ABA37573/c
 ID ABA37573 standard; DNA; 101 BP.

AC ABA37573;
 XX

DT 23-JAN-2002 (first entry)
 XX

DE Probe #16039 for gene expression analysis in human heart cell sample.
 XX

KW Human; gene expression; heart; microarray; vascular system; probe;
 XX cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.
 XX

OS Homo sapiens.
 XX

PN WO200157274-A2.
 XX

PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 XX

PR 26-MAY-2000; 2000US-0207456.
 XX

PR 30-JUN-2000; 2000US-0608408.
 XX

PR 03-AUG-2000; 2000US-0632366.
 XX

PR 21-SEP-2000; 2000US-0234687.
 XX

PR 27-SEP-2000; 2000US-0236359.
 XX

PR 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX

PT Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-488899/53.
 XX

PT Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts -

PS Claim 4; SEQ ID NO 16039; 530pp; English.
 XX

CC The present invention relates to single exon nucleic acid probes for
 XX measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for
 XX predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the
 XX probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system
 XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.
 XX Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
 XX

Query Match 3.7%; Score 101; DB 22; Length 101;
 XX Best Local Similarity 100.0%; Pred. No. 3.6e-39;

Db 533 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 592
 XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 593 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 592
 XX 101 AAAGATGTCTAATTCATCAAGTACTTAAGGCATTAATCAAGTCACTTGTACGC 42

QY 593 TGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 633
 Db 41 TGTGGAACGGGGGCTTTTCATCAATTTGCACTTGAATTTGACATCTCTGA 1

RESULT 39

AAK19593/c

```

ID  AAK19593 standard; DNA; 101 BP.
XX
AC  AAK19593;
XX
DT  05-NOV-2001 (first entry)
XX
DE  Human brain expressed single exon probe SEQ ID NO: 19584.
XX
KW  Human; brain expressed exon; gene expression analysis; probe;
KW  microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW  epilepsy; cancer; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157275-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00667.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-483446/52.
XX
PT  Single exon nucleic acid probes for analyzing gene expression in human
PT  brains -
XX
PS  Example 4; SEQ ID NO: 19584; 650bp + Sequence Listing; English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in the human
CC  brain. They can be used to measure gene expression in brain cell samples,
CC  which may enable the diagnosis and improved treatment of nervous system
CC  diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC  epilepsy and cancers. The present sequence is one of the probes of the
CC  invention.
XX
SQ  Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
XX
Query Match      3.7%; Score 101; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  533 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTATTAATCAAGCTGTTAGCGC 592
DB  101 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTATTAATCAAGCTGTTAGCGC 42

QY  593 TGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 633
DB  41 TGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 1

RESULT 40
ID  AAK45596/C
XX  AAK45596 standard; DNA; 101 BP.
XX
AC  AAK45596;
XX
DT  06-NOV-2001 (first entry)
XX
DE  Human bone marrow expressed single exon probe SEQ ID NO: 20153.
XX
KW  Human; bone marrow expressed exon; gene expression analysis; probe;
KW

```

```

KW  microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157276-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00668.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  30-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0632366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0236359.
PR  04-OCT-2000; 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI  Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR  WPI; 2001-488900/53.
XX
PT  Human genome-derived single exon nucleic acid probes useful for
PT  analyzing gene expression in human bone marrow -
XX
PS  Example 4; SEQ ID NO: 20153; 658bp + Sequence Listing; English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in the human
CC  bone marrow. They can be used to measure gene expression in bone marrow
CC  samples, which may enable the improved diagnosis and treatment of cancers
CC  such as lymphoma, leukemia and myeloma. The present sequence is one of
CC  the probes of the invention.
XX
SQ  Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
XX
Query Match      3.7%; Score 101; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  533 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTATTAATCAAGCTGTTAGCGC 592
DB  101 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTATTAATCAAGCTGTTAGCGC 42

QY  593 TGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 633
DB  41 TGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 1

RESULT 41
ID  AA151533/C
XX  AA151533 standard; DNA; 101 BP.
XX
AC  AA151533;
XX
DT  17-OCT-2001 (first entry)
XX
DE  Probe #20219 used to measure gene expression in human placenta sample.
XX
KW  Probe; microarray; human; placenta; antenatal diagnosis;
KW  genetic disorder; ss.
XX
OS  Homo sapiens.
XX
PN  WO200157272-A2.
XX
PD  09-AUG-2001.
XX
PF  30-JAN-2001; 2001WO-US00663.
XX
PR  04-FEB-2000; 2000US-0180312.

```



```

XX XX Human secreted protein 5' EST, SEQ ID NO: 2555.
DE PN
XX XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-0200610.
XX PR 26-FEB-1999; 99US-0122487.
XX PA (GSEST ) GENSEST.
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR P-PSDB; AAG02551.
XX PS WPI; 2000-500381/45.
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX XX
XX PS Claim 1, SEQ ID 2555; 71pp + CD-ROM; English.
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX CC well suited for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been
XX CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX CC mRNAs with intact 5' ends and can therefore be used to obtain full length
XX CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX CC gene therapy and chromosome mapping procedures. They are used to obtain
XX CC upstream regulatory sequences and to design expression and secretion
XX CC vectors.
XX SQ Sequence 229 BP; 63 A; 47 C; 53 G; 62 T; 4 other;
XX
XX Query Match 2.3%; Score 63; DB 21; Length 229;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-20;
XX Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 174 GCAAGGGAAGCTTACGACGCTGACATGAGTGTGCTTTCT 233
DB |||||
DB 123 GCAAGGGAAGCTTACGACGCTGACATGAGTGTGCTTTCT 182
QY 234 GGG 236
DB |||
DB 183 GGG 185
XX
XX RESULT 46
XX ABN43098
XX ID ABN43098 standard; DNA; 60 BP.
XX AC
XX ABN43098;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:15846.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS

```

```

XX XX WO200210449-A2.
DE PN
XX XX 07-FEB-2002.
XX PD
XX PF 20-JUL-2001; 2001WO-IB01903.
XX PR 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR WPI; 2002-257383/30.
XX
XX PT New oligonucleotide libraries comprising oligonucleotides which
XX PT selectively hybridize to mRNAs transcribed from a transcription unit of
XX PT a genome, useful for detecting tissue-, pathology-, and
XX PT developmental-specific genes -
XX PS Example 1, SEQ ID 15846; 47pp; English.
XX
XX CC The present invention describes oligonucleotide libraries for detecting
XX CC messenger RNAs that populate a (sub-)transcriptome, where the
XX CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX CC transcription units that populate a genome. The library comprises
XX CC several oligonucleotides, each capable of hybridizing selectively to a
XX CC set of messenger RNAs transcribed from a given transcription unit of
XX CC the genome, which encodes one or more messenger RNA splice variants.
XX CC The oligonucleotide libraries are useful for detecting mRNAs from a
XX CC biological sample, in expression profiling studies, in qualitatively or
XX CC quantitatively characterizing the corresponding transcriptome, and in
XX CC detecting RNA transcripts and splice variants of human or animal
XX CC transcripts. The libraries may also be used as specialised mini
XX CC libraries to detect transcripts of a sub-transcriptome under a
XX CC particular biological or pathological state, and so allowing the
XX CC detection of tissue- and pathology-specific genes such as those genes
XX CC only expressed in specific tissue under a specific pathological
XX CC condition, to detect developmental specific genes, and to detect RNA
XX CC transcripts and splice variants of a transcriptome of a patient suffering
XX CC from a particular disorder. ABN27253 to ABN59589 represent
XX CC oligonucleotide sequences from rats, humans and mice, which are used in
XX CC the exemplification of the present invention.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIRO
XX CC at ftp.wipo.int/pub/published_poc_sequences.
XX SQ Sequence 60 BP; 23 A; 11 C; 15 G; 11 T; 0 other;
XX
XX Query Match 2.2%; Score 60; DB 24; Length 60;
XX Best Local Similarity 100.0%; Pred. No. 6.4e-19;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 950 GCAATAGATGAGAACTCTGGAAGCTTACGCTAGATGATGATGCAAG 1009
DB |||||
DB 1 GCAATAGATGAGAACTCTGGAAGCTTACGCTAGATGATGATGCAAG 60
XX
XX RESULT 47
XX ABZ04740
XX ID ABZ04740 standard; DNA; 50 BP.
XX AC
XX ABZ04740;
XX
XX 09-JAN-2003 (first entry)
XX
XX Human leukocyte gene expression profiling probe SEQ ID NO 4731.
XX
XX T7; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection;
KW probe; ss.

```

```

XX OS Homo sapiens.
XX XX WO200257414-A2.
XX PN 25-JUL-2002.
XX PD
XX PF 22-OCT-2001; 2001WO-US47856.
XX XX
XX PR 20-OCT-2000; 2000US-241994P.
XX PR 08-JUN-2001; 2001US-296764P.
XX PA
XX PA (BIOC-) BIOCARDIA INC.
XX PI Wohlgenuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J,
XX PI Ly N, Woodward R, Querehemous T, Johnson F;
XX DR WPI; 2002-636525/68.
XX XX
XX PT New system for leukocyte expression profiling, diagnosing a disease, or
XX PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
XX PT or congestive heart failure, comprises diagnostic oligonucleotides -
XX XX
XX PS Claim 1, Page 479; 2038pp; English.
XX XX
XX CC The invention relates to a system for detecting gene expression, which
XX CC comprises one or two isolated DNA molecules that detect expression of a
XX CC gene, where the gene corresponds to any of 8143 oligonucleotides
XX CC (AB000010-AB208152) each having 50 base pairs (bp). The system is useful
XX CC for leukocyte expression profiling. It is particularly useful for
XX CC diagnosing a disease, monitoring (rate of) progression of a disease,
XX CC predicting therapeutic outcome, determining prognosis for a patient,
XX CC predicting disease complications in an individual or monitoring response
XX CC to treatment in an individual. The diseases include cardiac allograft
XX CC rejection, kidney allograft rejection, liver allograft rejection,
XX CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
XX CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection.
XX SQ
XX Sequence 50 BP; 19 A; 13 C; 7 G; 11 T; 0 other;
XX
XX Query Match 1.8%; Score 50; DB 24; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 5.6e-14;
XX Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2516 GCTGATTACCTCTAGAAACCTCAACAGTAGAACTTGCTAGACAT 2565
DB 1 GCTGATTACCTCTAGAAACCTCAACAGTAGAACTTGCTAGACAT 50
RESULT 48
AAD28273/C
ID AAD28273 standard; DNA; 4661 BP.
XX
XX AAD28273;
XX
XX 22-APR-2002 (first entry)
XX
XX Alpha-lactalbumin (LA) cc49IL2 vector.
XX
XX Bovine; alpha-lactalbumin; promoter; pharmaceutical; industrial; human;
XX KM diagnostic; screening; moloney murine leukemia virus; MoMuLV; chimeric;
XX ds.
XX
XX Chimeric - Bos sp.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Moloney murine leukemia virus.
XX OS Chimeric - Unidentified.
XX
XX Key Location/Qualifiers
XX FH 1.2055
XX FT misc_feature /tag= a
XX FT /note= "Bovine/human alpha-lactalbumin 5' flanking
XX FT region"

```

```

FT CDS 2098..4011
FT FT /*tag= b
FT FT /product= "cc49-IL2 coding region"
FT LTR 4068..4661
FT FT /*tag= c
FT FT /note= "MoMuLV 3' LTR"
XX
XX WO200202738-A2.
XX PN
XX PD 10-JAN-2002.
XX XX
XX PF 29-JUN-2001; 2001WO-US20710.
XX XX
XX PR 03-JUL-2000; 2000US-215925P.
XX XX
XX PA (GALA-) GALA DESIGN INC.
XX XX
XX PI Bremel RD, Miller LU, Bleck GT, York D;
XX XX
XX DR WPI; 2002-154737/20.
XX XX
XX PT Host cell for producing a desired protein and for screening compounds
XX PT useful for pharmaceutical, industrial, diagnostic and other purposes,
XX PT comprises multiple integrating vectors having an exogenous gene -
XX XX
XX PS Example 1; Fig 13; 191pp; English.
XX XX
XX CC The invention relates to a host cell comprising a genome having at least
XX CC two integrated integrating vectors. The integrating vectors comprise at
XX CC least one exogenous gene operably linked to a promoter. The host cell
XX CC is useful for producing a desired protein and for comparing protein
XX CC functions. The host cell comprises a reporter gene which is from gene
XX CC fluorescent protein, luciferase, beta-galactosidase and beta-lactamase,
XX CC and the assaying step further comprises detecting a signal from the
XX CC reporter gene. The desired protein includes proteins for pharmaceutical,
XX CC industrial, diagnostic and other purposes. The host cell is useful for
XX CC indirectly detecting the expression of a desired protein, comprising
XX CC providing the host cell transfected with a vector encoding a
XX CC polycistronic sequence comprising a signal protein and a desired protein
XX CC operably linked by an internal ribosome entry site (IRES), and culturing
XX CC the host cell under suitable conditions so that the signal protein and
XX CC the desired protein is produced, where the presence of signal protein
XX CC indicates the presence of desired protein. The present sequence is
XX CC alpha lactalbumin (LA) cc49IL2 vector used in the invention. The vector
XX CC comprises the following elements: bovine/human alpha-lactalbumin hybrid
XX CC promoter, cc49-IL2 coding region and 3' moloney murine leukemia virus
XX CC (MoMuLV) LTR.
XX
XX SQ Sequence 4661 BP; 1257 A; 1096 C; 1161 G; 1147 T; 0 other;
XX
XX Query Match 1.1%; Score 31; DB 24; Length 4661;
XX Best Local Similarity 100.0%; Pred. No. 0.00012;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTATTTATCGATGTTAACAGCTT 31
DB 4041 AATCTTTATTTATCGATGTTAACAGCTT 4011
RESULT 49
AAD28312/C
ID AAD28312 standard; DNA; 4661 BP.
XX
XX AAD28312;
XX
XX 22-APR-2002 (first entry)
XX
XX Alpha-lactalbumin (LA) cc49IL2 vector.
XX
XX Regulatory element; vector; erythropoietin; growth hormone; insulin;
XX KM immunoglobulin; bone morphogenetic protein; interferon; interleukin;
XX KM superoxide dismutase; T-cell receptor; surface membrane protein;
XX KM viral antigen; transport protein; addressin; regulatory protein;

```

KW	bovine; alpha-lactalbumin; promoter; human; MoMuLV; chimeric;
KW	moloney murine leukemia virus; ds.
XX	
OS	Chimeric - Bos sp.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Moloney murine leukemia virus.
OS	Chimeric - Unidentified.
XX	
FH	Key
FT	misc_feature
FT	1..2035
FT	/tag= a
FT	/note= "Bovine/human alpha-lactalbumin 5' flanking
FT	region"
CDS	2038..4011
FT	/tag= b
FT	/product= "cc49-IL2 coding region"
FT	4068..4661
FT	/tag= c
FT	/note= "MoMuLV 3' LTR"
XX	
PN	M0200202783-A2.
XX	
PD	10-JAN-2002.
XX	
PF	29-JUN-2001; 2001WO-US20714.
PR	03-JUL-2000; 2000US-215851P.
PA	(GALTA-) GALA DESIGN INC.
PI	Bleek GT;
XX	
XX	WPI; 2002-154749/20.
PT	
PT	Novel regulatory elements including nucleic acid encoding hybrid
PT	alpha-lactalbumin promoter or mutant RNA export element, for expressing
PT	one or more proteins e.g. antibodies, pharmaceutical proteins in host
XX	cells -
PS	
XX	Example 1; Fig 13; 151p; English.
CC	The invention relates to novel regulatory elements and vectors for the
CC	expression of one or more proteins in a host cell. The invention further
CC	provides methods of indirectly detecting the expression of a protein of
CC	interest, comprising providing the host cell transfected with a vector
CC	encoding a polypeptidic sequence comprising a signal protein and a
CC	desired protein operably linked by an internal ribosome entry site
CC	(IRES), and culturing the host cell under suitable conditions so that
CC	the signal protein and the desired protein is produced, where the
CC	presence of signal protein indicates the presence of desired protein.
CC	Regulatory elements and vectors of the invention are useful for the
CC	expression of proteins of interest in a host cell. They are useful for
CC	producing an immunoglobulin (Ig), preferably secretory Ig. They are
CC	useful in the expression of one or more proteins such as erythropoietin,
CC	growth hormone, insulin, immunoglobulins, protein C, cytokines and their
CC	receptors, hormones, Von Willebrand's factor, lung surfactant, serum
CC	albumins, DNase, vascular endothelial growth factor, receptors for
CC	hormones or growth factors, rheumatoid factors, nerve growth factors,
CC	CD proteins, osteoinductive factors, immunotoxins, bone morphogenetic
CC	protein, interferons, colony stimulating factors, interleukins,
CC	superoxide dismutase, T-cell receptors, surface membrane proteins,
CC	viral antigens, transport proteins, addressing, regulatory proteins,
CC	antibodies, chimeric proteins and their fragments. The vectors are
CC	particularly useful for expressing G protein coupled receptors and other
CC	transmembrane proteins. The retroviral vectors are useful for expressing
CC	proteins in mammalian tissue culture host cells, including rat fibroblast
CC	cells, bovine kidney cells and human kidney cells. The present sequence
CC	is alpha lactalbumin (IA) cc49II2 vector used in the invention. The
CC	vector comprises the following elements: bovine/human alpha-lactalbumin
CC	hybrid promoter; cc49-II2 coding region and LTR sequence from
CC	3' moloney murine leukemia virus (MoMuLV).
XX	
XQ	Sequence 4661 BP; 1257 A; 1096 C; 1161 G; 1147 T; 0 other;

OY	I AATCTTTATTATTTCGATGTAAACAAGCTT 31
ID	
ABZ25421	4041 AATCTTTATTATTTCGATGTAAACAAGCTT 4011
XX	
AC	ABZ25421;
XX	
DT	26-MAR-2003 (first entry)
XX	
DE	Human semaphorin-3A PCR primer #5.
XX	
KW	Human; PCR; primer; cytosolic; tumour; semaphorin-3A;
XX	prostatic cancer; BS.
OS	Homo sapiens.
XX	
PN	MO200286157-A2.
XX	
PD	31-OCT-2002.
XX	
PF	11-JAN-2002; 2002WO-FR00118.
XX	
PR	19-APR-2001; 2001US-284508P.
XX	
PA	(UROG-) UROGENE SA.
XX	
E1	Lacil A, Guesnot O, Algarre-Genin M;
XX	
DR	WPI; 2003-093153/08.
XX	
PT	In vitro evaluation of invasiveness of tumor cells, useful for
FT	diagnosis and monitoring of prostatic cancer, comprises measuring
XX	expression of semaphorin-3A -
PS	Disclosure; Page 17; 31pp; French.
XX	
CC	The present invention relates to a method for in vitro evaluation of the
CC	aggressiveness of tumor cells in a test sample. The method comprises
CC	measuring semaphorin-3A gene expression in both healthy and tumour
CC	epithelial cells. Cells are assessed as strongly invasive if the
CC	semaphorin-3A gene is underexpressed in them. The method is useful for
CC	diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC	assess the efficacy of treatments. The present sequence is a PCR primer
CC	for semaphorin-3A, used in the method of the invention.
XX	
XQ	Sequence 29 BP; 9 A; 6 C; 5 G; 9 T; 0 other;
XX	
Query Match	1.1%; Score 29; DB 25; Length 29;
Best Local Similarity	100.0%; Pred. No. 0.0013;
Matches	29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	2155 GGACATGGGTTTCATACAAACTCTTCTTA 2183
ID	
Dn	1 GGACATGGGTTTCATACAAACTCTTCTTA 29
XX	
RESULT 51	
AAF83093/C	AAF83093 standard; DNA; 9729 BP.
XX	
AC	AAF83093;
XX	
DT	29-JUN-2001 (first entry)
XX	

DE Nucleotide sequence of MLV self-inactivating (SIN) vector pTRAP2.
 XX
 XX Retrovirus; recombinase recognition sequence; RRS; LTR; recombinase;
 KM long terminal repeat; pharmacological; cytostatic; antiinflammatory;
 KM antirheumatic; antiallergic; antispasmodic; osteopathic; cardiac; MLV;
 KM vasotropic; neuroprotective; nootropic; cerebroprotective; antipsychotic;
 KM antileukemic; antileukemic; antileukemic; antileukemic; antileukemic;
 KM dermatological; gene therapy; ss.
 XX
 OS Synthetic.
 OS Murine leukemia virus.
 XX
 PN WO200125466-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-GB03837.
 XX
 PR 05-OCT-1999; 99GB-0023558.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Slingsby J, Kingsman SM, Rohli J, Slade A;
 DR WPI; 2001-281732/29.
 XX
 PT Modifying producer cells for making retrovirus by transfecting with a
 PT construct comprising a 5'- recombinase recognition sequence, long
 PT terminal repeat and 3'- recombinase recognition sequence, in presence
 PT of recombinase -
 PS
 PS Example 6; Page 123-126; 133pp; English.
 XX
 CC The invention relates to a method of modifying producer cells for making
 CC retrovirus by transfecting with a construct comprising a 5'- recombinase
 CC recognition sequence (RRS), long terminal repeat (LTR) and 3'- RRS, in
 CC presence of recombinase. The regulated retroviral vector produced is
 CC useful in the manufacture of a pharmaceutical composition to deliver a
 CC NOI to a target site, and in the manufacture of a medicament for
 CC diagnostic, therapeutic and/or medical applications. The recombinase
 CC assisted method is useful for introducing regulated 3'-LTR into a derived
 CC producer cell line to produce a high titer regulated retroviral vector.
 CC The vector is useful in gene therapy for treating diseases like cancers,
 CC inflammatory diseases, immunological disorders such as graft vs host
 CC disease, autoimmune diseases such as rheumatoid arthritis, allergic
 CC diseases such as asthma, osteoporosis, cardiovascular diseases such as
 CC congestive heart failure and ischemic heart disease, neurodegenerative
 CC disorders such as multiple sclerosis, Alzheimer's disease, stroke and
 CC cerebral ischemia, atherosclerosis, thrombotic disorders, dermatological
 CC disorders such as atopic dermatitis, contact dermatitis and psoriasis,
 CC wound healing, restenosis, infectious disorders such as HIV infections,
 CC ulcers, digestive disorders such as anorexia, bulimia and cachexia, and
 CC other diseases. The present sequence represents the nucleotide sequence
 CC of a Murine leukemia virus (MLV) self-inactivating (SIN) vector pTRAP2.
 XX
 SQ Sequence 9729 BP; 2270 A; 2657 C; 2540 G; 2262 T; 0 other;
 XX
 Query Match 1.0%; Score 27; DB 22; Length 9729;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTTAACAA 27
 |||||
 DB 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 52
 AAF83092/C
 ID AAF83092 standard; DNA; 12473 BP.
 XX
 AC .AAF83092;
 XX
 DT 29-JUN-2001 (first entry)

XX
 DE Nucleotide sequence of MLV construct CZCG.
 XX
 XX Retrovirus; recombinase recognition sequence; RRS; LTR; recombinase;
 KM long terminal repeat; pharmacological; cytostatic; antiinflammatory;
 KM antirheumatic; antiallergic; antispasmodic; osteopathic; cardiac; MLV;
 KM vasotropic; neuroprotective; nootropic; cerebroprotective; antipsychotic;
 KM antileukemic; antileukemic; antileukemic; antileukemic; antileukemic;
 KM dermatological; gene therapy; ss.
 XX
 OS Synthetic.
 OS Murine leukemia virus.
 XX
 PN WO200125466-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 05-OCT-2000; 2000WO-GB03837.
 XX
 PR 05-OCT-1999; 99GB-0023558.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Slingsby J, Kingsman SM, Rohli J, Slade A;
 DR WPI; 2001-281732/29.
 XX
 PT Modifying producer cells for making retrovirus by transfecting with a
 PT construct comprising a 5'- recombinase recognition sequence, long
 PT terminal repeat and 3'- recombinase recognition sequence, in presence
 PT of recombinase -
 PS
 PS Examples; Page 120-123; 133pp; English.
 XX
 CC The invention relates to a method of modifying producer cells for making
 CC retrovirus by transfecting with a construct comprising a 5'- recombinase
 CC recognition sequence (RRS), long terminal repeat (LTR) and 3'- RRS, in
 CC presence of recombinase. The regulated retroviral vector produced is
 CC useful in the manufacture of a pharmaceutical composition to deliver a
 CC NOI to a target site, and in the manufacture of a medicament for
 CC diagnostic, therapeutic and/or medical applications. The recombinase
 CC assisted method is useful for introducing regulated 3'-LTR into a derived
 CC producer cell line to produce a high titer regulated retroviral vector.
 CC The vector is useful in gene therapy for treating diseases like cancers,
 CC inflammatory diseases, immunological disorders such as graft vs host
 CC disease, autoimmune diseases such as rheumatoid arthritis, allergic
 CC diseases such as asthma, osteoporosis, cardiovascular diseases such as
 CC congestive heart failure and ischemic heart disease, neurodegenerative
 CC disorders such as multiple sclerosis, Alzheimer's disease, stroke and
 CC cerebral ischemia, atherosclerosis, thrombotic disorders, dermatological
 CC disorders such as atopic dermatitis, contact dermatitis and psoriasis,
 CC wound healing, restenosis, infectious disorders such as HIV infections,
 CC ulcers, digestive disorders such as anorexia, bulimia and cachexia, and
 CC other diseases. The present sequence represents the nucleotide sequence
 CC of a Murine leukemia virus (MLV) construct CZCG.
 XX
 SQ Sequence 12473 BP; 2841 A; 3434 C; 3357 G; 2841 T; 0 other;
 XX

OY 1 AATCTTTATTTATCGATGTTAACAA 27
 |||||
 DB 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 53
 AAK52430/C
 ID AAK52430 standard; CDNA; 1313 BP.
 XX
 AC .AAK52430;
 XX

DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 975.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QM, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM79297.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 3186-3187; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM79297-AAK80302) that exhibit actively elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1313 BP; 341 A; 304 C; 241 G; 427 T; 0 other;
XX
Query Match 1.0%; Score 26; DB 22; Length 1313;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 GTGTGACCAACGCGTCCGGAGTAG 80
DB 1305 GTGTGACCAACGCGTCCGGAGTAG 1280

RESULT 54
AB225416/c
ID AB225416 standard; DNA; 25 BP.
XX
AC AB225416;
XX
DT 26-MAR-2003 (first entry)

XX
DE Human semaphorin-3A PCR primer #2.
XX
KW Human; PCR; primer; cytoskeletal; tumour; semaphorin-3A;
KW prostatic cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200286157-A2.
XX
PD 31-OCT-2002.
XX
PF 11-JAN-2002; 2002WO-FR00118.
XX
PR 19-APR-2001; 2001US-284508P.
XX
PA (UROG-) UROGENE SA.
XX
PI Latil A, Cussenot O, Algarte-Genin M;
XX
DR WPI; 2003-093153/08.
XX
PT In vitro evaluation of invasiveness of tumor cells, useful for
PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
PS Disclosure; Page 16; 31pp; French.
XX
CC The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumor cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
SQ Sequence 25 BP; 7 A; 7 C; 4 G; 7 T; 0 other;
XX
Query Match 0.9%; Score 25; DB 25; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1556 ACAGATGTGGACCGTCTTAAG 1580
DB 25 ACAGATGTGGACCGTCTTAAG 1

RESULT 55
AAH47052
ID AAH47052 standard; DNA; 1399 BP.
XX
AC AAH47052;
XX
DT 29-OCT-2001 (first entry)
XX
DE Murine maspin cDNA sequence.
XX
KW Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;
KW 24p3; proliferin; maspin; cancer; cytoskeletal; gene therapy; ss.
XX
OS Mus sp.
XX
PN WO200155455-A2.
XX
PD 02-AUG-2001.
XX
PF 31-JAN-2001; 2001WO-US03161.
XX
PR 31-JAN-2000; 2000US-0179191.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
DT (JINS/) JIN S.

XX Jin S;
PI
XX
XX
DR WPI; 2001-488799/53.
XX
PT Determining if a compound modulates the drug resistance of a cell,
PT comprises determining the expression or activity level of a resistance
PT sequence in a cell in the presence of the test compound -
XX
XX Example 1; Fig 4; 79pp; English.
XX
XX The invention relates to a method of determining whether a test compound
CC modulates the drug resistance of a cell that comprises determining the
CC expression or activity level of resistance genes (e.g. semaphorin D, B94,
CC mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
CC the test compound, and comparing its expression or activity level in a
CC cell without the test compound. The drug resistance sequences are useful
CC in identifying drug resistant cells, in screening methods directed to the
CC identification of compounds that can modulate the drug resistance of a
CC cell type or multiple cell types. An isolated resistance protein can be
CC used as an immunogen to generate antibodies that bind the resistance
CC protein. Resistance nucleic acids may be inserted into vectors and used
CC as gene therapy vectors. An anti-resistance protein antibody may be used
CC to isolate a resistance protein from cells and of recombinantly produced
CC natural resistance protein from cells and of recombinantly produced
CC resistance protein expressed in host cells. The methods are useful for
CC treating a subject having a disorder, such as a drug-resistance cancer,
CC characterized by aberrant resistance sequence expression or activity by
CC administering to the subject a resistance sequence modulator. The present sequence
CC represents a murine maspin cDNA sequence, whose expression was decreased
CC in drug resistant EMT6 tumours.
XX
SQ Sequence 1399 BP; 436 A; 313 C; 278 G; 372 T; 0 other;
XX
Query Match 0.9%; Score 25; DB 22; Length 1399;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 22 TAACAAGCTTAGTAATGATGCCAC 46
DB 29 TAACAAGCTTAGTAATGATGCCAC 53
XX
RESULT 56
AB225415
ID AB225415 strand; DNA; 24 BP.
XX
XX AB225415;
AC
XX
XX 26-MAR-2003 (first entry)
DT
XX
XX Human semaphorin-3A PCR primer #1.
DE
XX
XX Human; PCR; primer; cytostatic; tumour; semaphorin-3A;
KM prostatic cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200286157-A2.
PN
XX
XX 31-OCT-2002.
PD
XX
XX 11-JAN-2002; 2002WO-FR00118.
PF
XX
XX 19-APR-2001; 2001US-284508P.
PR
XX
XX (UROG-) UROGENE SA.
PA
XX
XX Latil A, Cussenot O, Algarde-Genin M;
PI
XX
XX WPI; 2803-093153/08.
DR
XX
PT In vitro evaluation of invasiveness of tumor cells, useful for

PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
XX
XX Disclosure; Page 16; 31pp; French.
PS
XX
XX The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumour cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
SQ Sequence 24 BP; 8 A; 7 C; 4 G; 5 T; 0 other;
XX
Query Match 0.9%; Score 24; DB 25; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1442 CCTATGAACAATGCCCAATAGTG 1465
DB 1 CCTATGAACAATGCCCAATAGTG 24
XX
RESULT 57
AAC66819/C
ID AAC66819 standard; DNA; 8417 BP.
XX
XX AAC66819;
AC
XX
XX 27-FEB-2001 (first entry)
DT
XX
XX Reporter gene construct vector pRTRapsin-TK/BSN3.5'.
DE
XX
XX Reporter gene construct; gene trap; gene function; biological pathway;
KM
XX
XX Isotype switching; drug identification; ds.
KW
XX
XX Synthetic.
OS
XX
XX WO200061809-A2.
PN
XX
XX 19-OCT-2000.
PD
XX
XX 07-APR-2000; 2000WO-US09267.
PF
XX
XX 09-APR-1999; 99US-0128631.
PR
XX
XX (ICON-) ICONIX PHARM INC.
PA
XX
XX Xu H, Jarnigan K, Zhou H, Greene A, Thode S;
PI
XX
XX WPI; 2000-679501/66.
DR
XX
XX Reporter constructs for assaying biological response modifiers,
PT comprises a splice acceptor sequence, a first element comprising three
PT subelements, a reporter sequence, and a positive and negative
PT selectable marker -
XX
XX
XX Disclosure; Page 35-37; 57pp; English.
PS
XX
XX The present invention relates to novel reporter gene constructs,
CC designated gene traps, comprising a splice acceptor sequence, a reporter
CC sequence and positive and negative selectable markers. They can be used
CC in the expression of heterologous genes, to determine gene function, in
CC biological pathway analysis, to understand the mechanisms of isotype
CC switching and to identify drugs, particularly allergy drugs which
CC interfere with IgE production. The present sequence is one of the
CC reporter gene constructs of the invention.
XX
SQ Sequence 8417 BP; 1892 A; 2354 C; 2199 G; 1972 T; 0 other;
XX
Query Match 0.9%; Score 24; DB 21; Length 8417;

Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AATCTTTATTTATCGATGTTAA 24
Db 5703 AATCTTTATTTATCGATGTTAA 5680

RESULT 58
AAC66820/c
ID AAC66820 standard; DNA; 8684 BP.
XX
AC AAC66820;
XX
DT 27-FEB-2001 (first entry)
XX
DE Reporter gene construct vector pRTrip-TK/BSN3'5'.
XX
XX Reporter gene construct; gene trap; gene function; biological pathway;
KW isotype switching; drug identification; ds.
XX
OS Synthetic.
XX
PM WO200061809-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09267.
XX
PR 09-APR-1999; 99US-0128631.
XX
PA (ICON-) ICONIX PHARM INC.
XX
PI Xu H, Jarnigan K, Zhou H, Greene A, Thode S;
XX
DR WPI; 2000-679501/66.
XX
PT Reporter constructs for assaying biological response modifiers,
PT comprises a splice acceptor sequence, a first element comprising three
PT subelements, a reporter sequence, and a positive and negative
PT selectable marker -
XX
XX Disclosure; Page 38-40; 57pp; English.
XX
PS The present invention relates to novel reporter gene constructs,
CC designated gene traps, comprising a splice acceptor sequence, a reporter
CC sequence and positive and negative selectable markers. They can be used
CC in the expression of heterologous genes, to determine gene function, in
CC biological pathway analysis, to understand the mechanisms of isotype
CC switching and to identify drugs, particularly allergy drugs which
CC interfere with IgE production. The present sequence is one of the
CC reporter gene constructs of the invention.
XX
SQ Sequence 8684 BP; 1975 A; 2415 C; 2271 G; 2023 T; 0 other;
Query Match 0.9%; Score 24; DB 21; Length 8684;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AATCTTTATTTATCGATGTTAA 24
Db 5703 AATCTTTATTTATCGATGTTAA 5680

RESULT 59
AB225419
ID AB225419 standard; DNA; 29 BP.
XX
AC AB225419;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human semaphorin-3A PCR primer #3.

XX
XX Human; PCR; primer; cytoskeletal; tumour; semaphorin-3A;
KW prostatic cancer; ss.
XX
OS Homo sapiens.
XX
PM WO200286157-A2.
XX
PD 31-OCT-2002.
XX
PF 11-JAN-2002; 2002WO-FR00118.
XX
PR 19-APR-2001; 2001US-284508P.
XX
PA (UROG-) UROGENE SA.
XX
PI Latil A, Cussenot O, Algarre-Genin M;
XX
DR WPI; 2003-093153/08.
XX

PT In vitro evaluation of invasiveness of tumor cells, useful for
PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
XX Disclosure; Page 17; 31pp; French.

XX
XX The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumor cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
SQ Sequence 29 BP; 6 A; 6 C; 9 G; 8 T; 0 other;

Query Match 0.8%; Score 23; DB 25; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 192 TCTGACATGGGCTGTTAACT 214
Db 7 TCTGACATGGGCTGTTAACT 29

RESULT 60
ABK87383/c
ID ABK87383 standard; CDNA; 175 BP.
XX
AC ABK87383;
XX
DT 24-SEP-2002 (first entry)
XX
DE CDNA fragment for pertubagen R2.

XX
XX Retinoic acid pathway; RA; ss; retinoid; lung cancer;
KW Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
KW Darier's disease; Reiter's disease; psoriasis; gene therapy; R2;
KW acute promyelocytic leukaemia; APL; Pertubagen; R3; F802; F820;
KW yeast two-hybrid assay; PAT1; kinesin light chain-related protein;
KW human; EGFP; green fluorescent protein.
XX
OS Homo sapiens.
OS Aequorea victoria.
OS Synthetic.
XX
PM WO200240719-A2.
XX
PD 23-MAY-2002.
XX
PF 17-NOV-2001; 2001WO-US44039.

```

XX 17-NOV-2000; 2000US-249468P.
XX (DELTA-) DELTAGEN PROTEOMICS INC.
XX Kamb CA, Richards BT, Karpilow J;
XX WPI; 2002-519386/55.
XX
XX Polypeptide with retinoic acid pathway activity, especially of
XX perturbagens R3, F802 and F820 for identifying a cellular target which
XX interacts with the polypeptide and for therapeutic purposes
XX
XX Example 4; Fig 15a; 131pp; English.
XX
XX The invention relates to an isolated polypeptide (I) with retinoic acid
XX (RA) pathway activity, comprising a polypeptide sequence of Perturbagen
XX (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
XX biologically active modifications, or biologically active fragments.
XX Also included are the polynucleotides encoding the perturbagens, a
XX vector comprising the polynucleotide, preparing an RA pathway related
XX polypeptide, a composition comprising the polypeptide, an antibody to the
XX polypeptide, screening (M) putative RA-related therapeutics, by exposing
XX a polypeptide/target interaction pair identified by a method utilizing
XX the RA pathway polypeptide, to a number of agents, and recovering a
XX subpopulation of disrupting agents which competitively displace the
XX polypeptide from the target, where the disrupting agents are putative RA-
XX related therapeutics. Also include are an isolated RA pathway polypeptide
XX comprising PARI (a kinesin light chain-related protein) polypeptide
XX and its encoding polynucleotide, a gene therapy vector comprising the RA
XX pathway protein polypeptide or encoding or PARI and a host cell
XX comprising the gene therapy vector. The RA pathway polypeptide is useful
XX for identifying a cellular target that interacts with RA pathway-related
XX polypeptide, by exposing the polypeptide in vitro to putative target
XX molecules and identifying a polypeptide/target interaction pair, by
XX detecting reporter expression, where the reporter expression is
XX operatively linked to the formation of the interaction pair. (M) is a
XX yeast two-hybrid assay. The polypeptide is also useful for treating
XX an RA pathway-related condition e.g. lung cancer, Karpov's sarcoma,
XX breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
XX ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
XX Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
XX The present sequence is a cDNA fragment representing perturbagen R2
XX and containing sequences derived from the vector (EGFP, inactive green
XX fluorescent protein sequence).
XX
XX Sequence 175 BP; 40 A; 34 C; 60 G; 41 T; 0 other;
XX
XX Query Match 0.8%; Score 23; DB 24; Length 175;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AATCTTTATTTATCGATGTTA 23
XX |||||||||||||||||||
XX 169 AATCTTTATTTATCGATGTTA 147
XX
XX RESULT 61
XX ABR87385/c
XX ID ABR87385 standard; cDNA, 240 BP.
XX
XX AC ABR87385;
XX
XX 24-SEP-2002 (first entry)
XX
XX cDNA fragment for perturbagen R44.
XX
XX Retinoic acid pathway; RA; ss; retinoid; lung cancer;
XX Karpov's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
XX renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
XX Darier's disease; Reiter's disease; psoriasis; gene therapy; R44;
XX acute promyelocytic leukaemia; APL; Perturbagen; R3; F802; F820;
XX yeast two-hybrid assay; PARI; kinesin light chain-related protein;

```

```

XX human; EGFP, green fluorescent protein.
XX Homo sapiens.
XX Aequorea victoria.
XX Synthetic.
XX WO200240719-A2.
XX
XX 23-MAY-2002.
XX
XX 17-NOV-2001; 2001WO-US44039.
XX
XX 17-NOV-2000; 2000US-249468P.
XX
XX (DELTA-) DELTAGEN PROTEOMICS INC.
XX
XX Kamb CA, Richards BT, Karpilow J;
XX WPI; 2002-519386/55.
XX
XX Polypeptide with retinoic acid pathway activity, especially of
XX perturbagens R3, F802 and F820 for identifying a cellular target which
XX interacts with the polypeptide and for therapeutic purposes
XX
XX Example 4; Fig 15b; 131pp; English.
XX
XX The invention relates to an isolated polypeptide (I) with retinoic acid
XX (RA) pathway activity, comprising a polypeptide sequence of Perturbagen
XX (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
XX biologically active modifications, or biologically active fragments.
XX Also included are the polynucleotides encoding the perturbagens, a
XX vector comprising the polynucleotide, preparing an RA pathway related
XX polypeptide, a composition comprising the polypeptide, an antibody to the
XX polypeptide, screening (M) putative RA-related therapeutics, by exposing
XX a polypeptide/target interaction pair identified by a method utilizing
XX the RA pathway polypeptide, to a number of agents, and recovering a
XX subpopulation of disrupting agents which competitively displace the
XX polypeptide from the target, where the disrupting agents are putative RA-
XX related therapeutics. Also include are an isolated RA pathway polypeptide
XX comprising PARI (a kinesin light chain-related protein) polypeptide
XX and its encoding polynucleotide, a gene therapy vector comprising the RA
XX pathway protein polypeptide or encoding or PARI and a host cell
XX comprising the gene therapy vector. The RA pathway polypeptide is useful
XX for identifying a cellular target that interacts with RA pathway-related
XX polypeptide, by exposing the polypeptide in vitro to putative target
XX molecules and identifying a polypeptide/target interaction pair, by
XX detecting reporter expression, where the reporter expression is
XX operatively linked to the formation of the interaction pair. (M) is a
XX yeast two-hybrid assay. The polypeptide is also useful for treating
XX an RA pathway-related condition e.g. lung cancer, Karpov's sarcoma,
XX breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
XX ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
XX Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
XX The present sequence is a cDNA fragment representing perturbagen R44
XX and containing sequences derived from the vector (EGFP, inactive green
XX fluorescent protein sequence).
XX
XX Sequence 240 BP; 65 A; 32 C; 87 G; 56 T; 0 other;
XX
XX Query Match 0.8%; Score 23; DB 24; Length 240;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AATCTTTATTTATCGATGTTA 23
XX |||||||||||||||||||
XX 234 AATCTTTATTTATCGATGTTA 212
XX
XX RESULT 62
XX ABR87379/c
XX ID ABR87379 standard; cDNA, 287 BP.
XX
XX AC ABR87379;

```

24-SEP-2002 (first entry)
 CDNA encoding partial perturbation F802.
 Retinoic acid pathway; RA; ss; retinoid; lung cancer;
 Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
 renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
 Darier's disease; Reiter's disease; psoriasis; gene therapy; F797;
 acute promyelocytic leukemia; APL; Perturbation; R3; F802; F820;
 yeast two-hybrid assay; PARI; kinesin light chain-related protein;
 human; EGFP; green fluorescent protein.
 Homo sapiens.
 Aegorea victoria.
 Synthetic.
 Location/Qualifiers
 Key CDS
 /tag= a
 /product= "Partial perturbation F802"
 /partial
 /note= "No start codon shown"
 WO200240719-A2.
 23-MAY-2002.
 17-NOV-2001; 2001WO-US44039.
 17-NOV-2000; 2000US-249468P.
 (DELT-) DELTAGEN PROTEOMICS INC.
 Kamb CA, Richards BT, Karpilow J,
 WPI; 2002-519386/55.
 P-PSDB; AAU99181.
 Polypeptide with retinoic acid pathway activity, especially of
 perturbagens R3, F802 and F820 for identifying a cellular target which
 interacts with the polypeptide and for therapeutic purposes
 Claim 67; Fig 13a; 131pp; English.
 The invention relates to an isolated polypeptide (I) with retinoic acid
 (RA) pathway activity, comprising a polypeptide sequence of Perturbation
 (phenotypic probe) R3, Perturbation F802, Perturbation F820, their
 biologically active modifications, or biologically active fragments.
 Also included are the polynucleotides encoding the perturbagens, a
 vector comprising the polynucleotide, preparing an RA pathway related
 polypeptide, a composition comprising the polypeptide, an antibody to the
 polypeptide, screening (M) putative RA-related therapeutics, by exposing
 a polypeptide/target interaction pair identified by a method utilising
 the RA pathway polypeptide, to a number of agents, and recovering a
 subpopulation of disrupting agents which competitively displace the
 polypeptide from the target, where the disrupting agents are putative RA-
 related therapeutics. Also include are an isolated RA pathway polypeptide
 comprising PARI (a kinesin light chain-related protein) polypeptide
 and its encoding polynucleotide, a gene therapy vector comprising the RA
 pathway protein polypeptide or encoding or PARI and a host cell
 comprising the gene therapy vector. The RA pathway polypeptide is useful
 for identifying a cellular target that interacts with RA pathway-related
 polypeptide, by exposing the polypeptide in vitro to putative target
 molecules and identifying a polypeptide/target interaction pair, by
 detecting reporter expression, where the reporter expression is
 operatively linked to the formation of the interaction pair. (M) is a
 yeast two-hybrid assay. The polypeptide is also useful for treating
 an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 Reiter's disease, psoriasis, acute promyelocytic leukemia (APL).
 The present sequence is a CDNA encoding partial perturbation F802

and containing sequences derived from the vector (EGFP, inactive green
 fluorescent protein sequence).
 Sequence 287 BP; 93 A; 61 C; 74 G; 59 T; 0 other;

Query Match 0.8%; Score 23; DB 24; Length 287;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 AATCTTTATTTATCGATGTTA 23
 Db 280 AATCTTTATTTATCGATGTTA 258

RESULT 63

AA18251/c
 ID AA18251 standard; CDNA; 326 BP.

AA18251;

07-DEC-2001 (first entry)

Human breast cancer expressed polynucleotide 10708.

Human; breast cancer; cell marker; cytostatic; ss.

Homo sapiens.

WO200151628-A2.

19-JUL-2001.

10-JAN-2001; 2001WO-US00798.

14-JAN-2000; 2000US-0176077.

14-MAR-2000; 2000US-0189167.

24-MAR-2000; 2000US-0192099.

29-MAR-2000; 2000US-0193480.

15-MAY-2000; 2000US-0205230.

09-JUN-2000; 2000US-0211315.

25-JUL-2000; 2000US-0220534.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lillie J, Xu Y, Wang Y, Steinmann K;

WPI; 2001-451856/48.

New peptide useful as a marker for the diagnosis of breast cancer

Claim 1; Page 1911; 3695pp; English.

The invention relates to human breast cancer expressed polynucleotides
 (AAU07544-AAU26789) and methods of assessing whether a patient is
 afflicted with breast cancer by examining the correlation between the
 expression of certain markers and the cancerous state of breast cells.
 The polynucleotides and encoded polypeptides are potential markers for
 detecting, diagnosing, monitoring, characterizing treating and
 potentially preventing breast cancer. The polynucleotides and encoded
 polypeptides are also useful for isolating compounds with cytostatic
 activity.

Sequence 326 BP; 73 A; 51 C; 79 G; 123 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 326;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 GGTGTGACCCACCGCTCCGCGGA 76

Db 294 GGTGTGACCCACCGCTCCGCGGA 272

ID	ABK87384	standard; cDNA; 335 BP.
XX	ABK87384/	
DT	24-SEP-2002	(first entry)
XX		
DE	cDNA fragment for perturbation R3.	
XX		
KW	Retinoic acid pathway; RA; ss; retinoid; lung cancer;	
KW	Kaprosi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;	
KW	renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;	
KW	Darier's disease; Reiter's disease; psoriasis; gene therapy; R3;	
KW	acute promyelocytic leukaemia; APL; Perturbation; R3; F802; F820;	
KW	yeast two-hybrid assay; PAT1; kinesin light chain-related protein;	
KW	human; EGFP; green fluorescent protein.	
XX		
OS	Homo sapiens.	
OS	Aequorea victoria.	
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	44..73
FT		/*tag= a
FT	/partial	/product= "Partial perturbation R3"
FT		/note= "No start codon shown"
XX		
XX		
PN	MO200240719-A2.	
PD		
XX	23-MAY-2002.	
XX		
PF	17-NOV-2001; 2001MO-US44039.	
PR		
XX	17-NOV-2000; 2000US-249468P.	
XX		
PA	(DELT-) DELTAGEN PROTEOMICS INC.	
PI	Kamb CA, Richards BT, Karpilow J;	
DR	WPI; 2002-519366/55.	
DR	P-PSDB; AAU99184.	
XX		
PT	Poly peptide, used in retinoic acid pathway activity, especially of	
PT	perturbagens R3, F802 and F820 for identifying a cellular target which	
PT	interacts with the poly peptide and for therapeutic purposes	
XX		
PS	Claim 67; Fig 15a; 131pp; English.	

Query Match	0.8%	Score 23	DB 24	Length 335
Best Local Similarity	100.0%	Pred. No. 1.2		
Matches 23	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Oy	1	AATCTTTATTTATCGATGTTA	23	
Db	329	AATCTTTATTTATCGATGTTA	307	
<p>RESULT 65</p> <p>ABK87380/c</p> <p>ID ABK87380 standard; cDNA; 339 BP.</p> <p>XX ABK87380;</p> <p>XX 24-SEP-2002 (first entry)</p> <p>DE cDNA encoding partial perturbagen F820.</p> <p>XX Retinoic acid pathway; RA; ss; retinoid; lung cancer;</p> <p>KW Karpov's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;</p> <p>KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;</p> <p>KW Darier's disease; Reiter's disease; psoriasis; gene therapy; F797;</p> <p>KW acute promyelocytic leukaemia; APL; Perturbagen; R3; F802; F820;</p> <p>KW yeast two-hybrid assay; P411; kinesin light chain-related protein;</p> <p>XX human; EGFP; green fluorescent protein.</p> <p>XX Homo sapiens.</p> <p>OS Aequorea victoria.</p> <p>OS Synthetic.</p> <p>XX Key</p> <p>FT CDS</p> <p>XX Location/Qualifiers</p> <p>XX 44..136</p> <p>XX /*tag= a</p> <p>XX /product= "Partial perturbagen F820"</p> <p>XX /partial</p> <p>XX /note= "No start codon shown"</p> <p>XX /transl_except= (pos:56..61,aa:11e)</p> <p>XX WO200240719-A2.</p> <p>XX 23-MAY-2002.</p> <p>XX 17-NOV-2001; 2001WO-US44039.</p> <p>XX 17-NOV-2000; 2000US-249468P.</p> <p>XX (DELTA-) DELTAGEN PROTEOMICS INC.</p> <p>XX Kamb CA, Richards BT, Karpilow J;</p> <p>XX WP1; 2002-519386/55.</p> <p>XX P-PSDB; AAU99182.</p> <p>XX Polypeptide with retinoic acid pathway activity, especially of</p> <p>XX perturbagens R3, F802 and F820 for identifying a cellular target which</p> <p>XX interacts with the polypeptide and for therapeutic purposes -</p> <p>XX Claim 67; Fig 13b; 131pp; English.</p> <p>XX The invention relates to an isolated polypeptide (I) with retinoic acid</p> <p>XX (RA) pathway activity, comprising a polypeptide sequence of Perturbagen</p> <p>XX (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their</p>				

biologically active modifications, or biologically active fragments. Also included are the polynucleotides encoding the perturbagens, a vector comprising the polynucleotide, preparing an RA pathway related polypeptide, a composition comprising the polypeptide, an antibody to the polypeptide, screening (M) putative RA-related therapeutics, by exposing a polypeptide/target interaction pair identified by a method utilizing the RA pathway polypeptide, to a number of agents, and recovering a subpopulation of disrupting agents which competitively displace the polypeptide from the target, where the disrupting agents are putative RA-related therapeutics. Also include are an isolated RA pathway polypeptide comprising PRT1 (a kinase light chain-related protein) polypeptide and its encoding polynucleotide, a gene therapy vector comprising the RA pathway protein polypeptide or encoding or PRT1 and a host cell comprising the gene therapy vector. The RA pathway polypeptide is useful for identifying a cellular target that interacts with RA pathway-related polypeptide, by exposing the polypeptide in vitro to putative target molecules and identifying a polypeptide/target interaction pair, by detecting reporter expression, where the reporter expression is operatively linked to the formation of the interaction pair. (M) is a yeast two-hybrid assay. The polypeptide is also useful for treating an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma, breast cancer, pancreatic cancer, neuroblastoma, renal cancer, ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease, Reiter's disease, psoriasis, acute promyelocytic leukemia (APL). The present sequence is a cDNA encoding partial perturbagen F820 and containing sequences derived from the vector (EGFP, inactive green fluorescent protein sequence).

Sequence 339 BP, 119 A, 57 C, 64 G, 99 T, 0 other;

Query Match 0.8%; Score 23; DB 24; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATCTTTATTTATCGATGTTA 23

332 AATCTTTATTTATCGATGTTA 310

RESULT 66

ABK87382/C

ABK87382 standard; cDNA, 352 BP.

ABK87382;

24-SEP-2002 (first entry)

CDNA fragment for perturbagen R1.

Retinoic acid pathway; RA; ss; retinoid; lung cancer; Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma; renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema; Darier's disease; Reiter's disease; psoriasis; gene therapy; R1; acute promyelocytic leukemia; APL; perturbagen; R3; F802; F820; yeast two-hybrid assay; PRT1; kinesin light chain-related protein; human; EGFP; green fluorescent protein.

Homo sapiens.

Aequorea victoria.

Synthetic.

WO200240719-AZ.

23-MAY-2002.

17-NOV-2001; 2001WO-US44039.

17-NOV-2000; 2000US-249468P.

(DELT-) DELTAGEN PROTEOMICS INC.

Kamb CA, Richards BT, Karpilow J;

WPI; 2002-519386/55.

Polypeptide with retinoic acid pathway activity, especially of perturbagens R3, F802 and F820 for identifying a cellular target which interacts with the polypeptide and for therapeutic purposes

Example 4; Fig 15a; 131p; English.

The invention relates to an isolated polypeptide (I) with retinoic acid (RA) pathway activity, comprising a polypeptide sequence of Perturbagen (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their biologically active modifications, or biologically active fragments. Also included are the polynucleotides encoding the perturbagens, a vector comprising the polynucleotide, preparing an RA pathway related polypeptide, a composition comprising the polypeptide, an antibody to the polypeptide, screening (M) putative RA-related therapeutics, by exposing a polypeptide/target interaction pair identified by a method utilizing the RA pathway polypeptide, to a number of agents, and recovering a subpopulation of disrupting agents which competitively displace the polypeptide from the target, where the disrupting agents are putative RA-related therapeutics. Also include are an isolated RA pathway polypeptide comprising PRT1 (a kinase light chain-related protein) polypeptide and its encoding polynucleotide, a gene therapy vector comprising the RA pathway protein polypeptide or encoding or PRT1 and a host cell comprising the gene therapy vector. The RA pathway polypeptide is useful for identifying a cellular target that interacts with RA pathway-related polypeptide, by exposing the polypeptide in vitro to putative target molecules and identifying a polypeptide/target interaction pair, by detecting reporter expression, where the reporter expression is operatively linked to the formation of the interaction pair. (M) is a yeast two-hybrid assay. The polypeptide is also useful for treating an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma, breast cancer, pancreatic cancer, neuroblastoma, renal cancer, ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease, Reiter's disease, psoriasis, acute promyelocytic leukemia (APL). The present sequence is a cDNA fragment representing perturbagen R1 and containing sequences derived from the vector (EGFP, inactive green fluorescent protein sequence).

Sequence 352 BP, 77 A, 66 C, 92 G, 117 T, 0 other;

Query Match 0.8%; Score 23; DB 24; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATCTTTATTTATCGATGTTA 23

346 AATCTTTATTTATCGATGTTA 324

RESULT 67

ABK87381/C

ABK87381 standard; cDNA, 527 BP.

ABK87381;

24-SEP-2002 (first entry)

CDNA encoding partial perturbagen H241.

Retinoic acid pathway; RA; ss; retinoid; lung cancer; Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma; renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema; Darier's disease; Reiter's disease; psoriasis; gene therapy; H241; acute promyelocytic leukemia; APL; perturbagen; R3; F802; F820; yeast two-hybrid assay; PRT1; kinesin light chain-related protein; human; EGFP; green fluorescent protein.

Homo sapiens.

Aequorea victoria.

Synthetic.

Location/Qualifiers

FT CDS 44..83
FT /*tag= a
FT /product= "Partial perturbation H241"
FT /partial
FT /note= "No start codon shown"
XX
XX WO200240719-A2.
XX
XX 23-MAY-2002.
XX
XX
XX 17-NOV-2001; 2001WO-US44039.
XX
XX 17-NOV-2000; 2000US-249468P.
XX
XX (DELTA-) DELTAGEN PROTEOMICS INC.
XX
XX Kamb CA, Richards BT, Karpilow J;
XX WPI; 2002-519386/55.
XX P-PSDB; AAU99183.
XX
XX
XX Polypeptide with retinoic acid pathway activity, especially of
PT perturbation R3, F802 and F820 for identifying a cellular target which
PT interacts with the polypeptide and for therapeutic purposes
XX
XX Example 3; Fig 13b; 131pp; English.
XX
XX The invention relates to an isolated polypeptide (I) with retinoic acid
CC (RA) pathway activity, comprising a polypeptide sequence of Perturbation
CC (phenotypic probe) R3, Perturbation F802, Perturbation F820, their
CC biologically active modifications, or biologically active fragments.
CC Also included are the polynucleotides encoding the perturbagens, a
CC vector comprising the polynucleotide, preparing an RA pathway related
CC polypeptide, a composition comprising the polypeptide, an antibody to the
CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
CC a polypeptide/target interaction pair identified by a method utilizing
CC the RA pathway polypeptide, to a number of agents, and recovering a
CC subpopulation of disrupting agents which competitively displace the
CC polypeptide from the target, where the disrupting agents are putative RA-
CC related therapeutics. Also include are an isolated RA pathway polypeptide
CC comprising PRT1 (a kinesin light chain-related protein) polypeptide
CC and its encoding polynucleotide, a gene therapy vector comprising the RA
CC pathway protein polypeptide or encoding or PRT1 and a host cell
CC comprising the gene therapy vector. The RA pathway polypeptide is useful
CC for identifying a cellular target that interacts with RA pathway-related
CC polypeptide, by exposing the polypeptide in vitro to putative target
CC molecules and identifying a polypeptide/target interaction pair, by
CC detecting reporter expression, where the reporter expression is
CC operatively linked to the formation of the interaction pair. (M) is a
CC yeast two-hybrid assay. The polypeptide is also useful for treating
CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
CC Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
CC The present sequence is a cDNA encoding partial perturbation H241
CC and containing sequences derived from the vector (BGFP, inactive green
CC fluorescent protein sequence).
XX
XX Sequence 527 BP; 187 A; 81 C; 130 G; 129 T; 0 other;
SQ

Query Match 0.8%; Score 23; DB 24; Length 527;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTTATTTATGATGTTA 23
Db 520 AATCTTTTATTTATGATGTTA 498

RESULT 68
AAK63284
ID AAK63284 standard; cDNA; 632 BP.
XX

AC AAK63284;
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8344.
DE
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW
XX
XX cytosolic; gene therapy; vaccine; metastasis; ss.
OS
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 14-AUG-2000; 2000US-0226279.
XX 18-AUG-2000; 2000US-0226881.
XX 22-AUG-2000; 2000US-0226886.
XX 22-AUG-2000; 2000US-0227182.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0228287.
XX 01-SEP-2000; 2000US-0228343.
XX 01-SEP-2000; 2000US-0228344.
XX 01-SEP-2000; 2000US-0228345.
XX 05-SEP-2000; 2000US-0228309.
XX 05-SEP-2000; 2000US-0228513.
XX 06-SEP-2000; 2000US-0228517.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.

```

PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0235327.
PR 29-SEP-2000; 2000US-0235367.
PR 29-SEP-2000; 2000US-0235368.
PR 29-SEP-2000; 2000US-0235369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246612.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0255719.
PR 06-DEC-2000; 2000US-0251479.

```

```

PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR XX
PR XX (HUMA-) HUMAN GENOME SCI INC.
PR PA
PR PI Rosen CA, Barash SC, Ruben SM;
DR WPI; 2001-483426/52.
DR P-PSDB; AAM90503.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
FT
XX
XX Claim 1; SEQ ID NO 8344; 3071bp + Sequence Listing; English.
XX
PS AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 632 BP; 202 A; 131 C; 124 G; 170 T; 5 other;

Query Match 0.8%; Score 23; DB 22; Length 632;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 57 GTCGACCCACGCGTCGGAGCTA 79
DB 72 GTCGACCCACGCGTCGGAGCTA 94

RESULT 69
ABK87378/c
ID ABK87378 standard; cDNA; 661 BP.
XX
AC ABR87378;
XX
DT 24-SEP-2002 (first entry)
XX
DB cDNA encoding partial perturbation F797.
XX
XX Retinoic acid pathway; RA; ss; retinoid; lung cancer;
KW Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
KW Darier's disease; Reiter's disease; psoriasis; gene therapy; F797;
KW acute promyelocytic leukaemia; APL; Perturbation; R3; F802; F820;
KW yeast two-hybrid assay; P471; kinesin light chain-related protein;
KW human; EGFP; green fluorescent protein.
XX
XX Homo sapiens.
OS Aegonorea victoria.
OS Synthetic.
XX
FH Key Location/Qualifiers

```


FT CDS 44.112
 FT /*tag= a
 FT /product= "Partial perturbation F797"
 FT /partial
 FT /note= "No start codon shown"
 FT /transl_except= (pos:95..97,aa:leu)
 PN WO200240719-A2.
 XX
 XX 23-MAY-2002.
 PD
 XX
 XX 17-NOV-2001; 2001WO-US44039.
 XX
 XX 17-NOV-2000; 2000US-249468P.
 PR
 XX (DELTA-) DELTAGEN PROTEOMICS INC.
 XX
 XX Kamb CA, Richards BT, Karpilow J;
 PI
 XX WPI; 2002-519386/55.
 DR P-PSDB; AAU99180.
 DR
 XX
 XX
 PT Polypeptide with retinoic acid pathway activity, especially of
 PT perturbation R3, F802 and F820 for identifying a cellular target which
 PT interacts with the polypeptide and for therapeutic purposes -
 PS
 PS Example 3; Fig 13a; 131pp; English.

CC The invention relates to an isolated polypeptide (I) with retinoic acid
 CC (RA) pathway activity, comprising a polypeptide sequence of Perturbation
 CC (phenotypic probe) R3, Perturbation F802, Perturbation F820, their
 CC biologically active modifications, or biologically active fragments.
 CC Also included are the polynucleotides encoding the perturbations, a
 CC vector comprising the polynucleotide, preparing an RA pathway related
 CC polypeptide, a composition comprising the polypeptide, an antibody to the
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
 CC a polypeptide/target interaction pair identified by a method utilizing
 CC the RA pathway polypeptide, to a number of agents, and recovering a
 CC subpopulation of disrupting agents which competitively displace the
 CC polypeptide from the target, where the disrupting agents are putative RA-
 CC related therapeutics. Also include are an isolated RA pathway polypeptide
 CC comprising PRT1 (a kinesin light chain-related protein) polypeptide
 CC and its encoding polynucleotide, a gene therapy vector comprising the RA
 CC pathway protein polypeptide or encoding or PRT1 and a host cell
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful
 CC for identifying a cellular target that interacts with RA pathway-related
 CC polypeptide, by exposing the polypeptide in vitro to putative target
 CC molecules and identifying a polypeptide/target interaction pair, by
 CC detecting reporter expression, where the reporter expression is
 CC operatively linked to the formation of the interaction pair. (M) is a
 CC yeast two-hybrid assay. The polypeptide is also useful for treating
 CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 CC Reiter's disease, psoriasis, acute promyelocytic leukemia (APL).
 CC The present sequence is a cDNA encoding partial perturbation F797
 CC and containing sequences derived from the vector (BGFP, inactive green
 CC fluorescent protein sequence).
 XX

SO Sequence 661 BP; 248 A; 87 C; 156 G; 170 T; 0 other;

Query Match 0.8%; Score 23; DB 24; Length 661;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTTA 23
 |||||
 DB 654 AATCTTTATTTATCGATGTTA 632

RESULT 70
 AAH36207
 ID AAH36207 standard; cDNA, 115 BP.

XX
 AC AAH36207;
 XX
 XX 03-SEP-2001 (first entry)
 DT
 XX Human colon cancer antigen encoding cDNA SEQ ID NO:3289.
 DE
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; 88.
 XX
 OS Homo sapiens.
 PN WO200122920-A2.
 XX
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US26524.
 PR
 XX 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI
 XX WPI; 2001-235357/24.
 DR P-PSDB; AAG76802.
 DR
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 PS
 PS Claim 1; Page 5085-5086; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytoskeletal activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SO Sequence 115 BP; 20 A; 33 C; 34 G; 21 T; 7 other;

Query Match 0.8%; Score 22; DB 22; Length 115;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACCGCTCCGGAGT 78
 |||||
 DB 61 GTCGACCCACCGCTCCGGAGT 82

RESULT 71
 AAK61215
 ID AAK61215 standard; cDNA; 370 BP.
 XX
 XX AAK61215;
 AC
 XX
 XX 06-NOV-2001 (first entry)
 DT
 XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6275.
 DE

Human; immune; hematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ss.
Homo sapiens.
WO200157182-A2.
09-AUG-2001.
17-JAN-2001; 2001WO-US01354.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
04-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
14-JUL-2000; 2000US-0217496.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225265.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226881.
22-AUG-2000; 2000US-0226886.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229511.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0233401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233064.
14-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
01-NOV-2000; 2000US-0241826.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249300.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251858.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
 XX 05-JAN-2001; 2001US-02559678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 DR P-PSDB; AAM88434.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PS
 XX Claim 1; SEQ ID NO 6275; 3071pp + Sequence Listing; English.
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK7694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 370 BP; 62 A; 90 C; 146 G; 67 T; 5 other;
 Query Match 0.8%; Score 22; DB 22; Length 370;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 57 GTGCAGCCACGCGTCGGGAGT 78
 Db 3 GTGCAGCCACGCGTCGGGAGT 24
 RESULT 72
 AA192960
 ID AA192960 standard; cDNA; 471 BP.
 XX
 AC AA192960;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13020.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 XX
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Tang YT, Liu C, Dermanac RT;
 XX
 DR WPI; 2001-514838/56.
 XX P-PSDB; AAO13029.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 PS
 XX Claim 1; SEQ ID NO 13020; 1399pp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 471 BP; 80 A; 151 C; 112 G; 123 T; 5 other;
 Query Match 0.8%; Score 22; DB 22; Length 471;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 57 GTGCAGCCACGCGTCGGGAGT 78
 Db 43 GTGCAGCCACGCGTCGGGAGT 64
 RESULT 73
 AB577002
 ID AB577002 standard; cDNA; 711 BP.
 XX
 AC AB577002;
 XX
 DT 12-DEC-2002 (first entry)
 XX
 DE Frog embryonic gene sequence Q925394.
 XX
 KW Frog; ss; embryonic development; developmental disorder;
 KW microarray; cell differentiation.
 XX
 OS Xenopus laevis.
 XX
 PN US2002081610-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 23-JUL-2001; 2001US-0910943.
 XX
 PR 21-JUL-2000; 2000US-219658P.
 XX
 XX (UYRQ) UNIV ROCKEFELLER.
 XX
 PA Hemmati-Briuanlou A, Altmann CR;
 XX
 PI WPI; 2002-626534/67.
 XX
 DR Nucleic acid array containing Xenopus embryonic nucleic acids is useful
 XX to identify genes involved in embryonic development, to identify
 XX different types of embryonic cells, and to diagnose developmental
 XX disorders -
 PS
 XX Claim 1; Page 272; 823pp; English.
 XX

CC The invention relates to a nucleic acid array, where each coordinate
CC contains a single nucleic acid species having one of 770 nucleotide
CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene
CC product, or its complement or hybridisable fragment of not less than
CC 20 contiguous nucleotides of one of those sequences. Also included are
CC detecting differential expression of embryonic genes, comprising:
CC (a) contacting a nucleic acid array comprising genes expressed in
CC embryonic but not mature cells with nucleic acids from sample and
CC control cells; and (b) detecting differential hybridisation of nucleic
CC acids from the sample cells relative to the control cells; and detecting
CC defects in development, comprising: (a) contacting nucleic acids from
CC test cells undergoing development with a nucleic acid array of gene
CC products known to play a fundamental role in the development process; and
CC (b) detecting a difference in expression of a fundamental gene in the
CC sample cells relative to a standard. The invention is useful to identify
CC genes involved in embryonic development and related processes such as
CC cell differentiation. This would be useful for diagnosing developmental
CC disorders and for identifying different types of embryonic cells.
CC The present sequence is one of the 770 Xenopus embryonic cDNA sequences.
XX

SO Sequence 711 BP; 219 A; 173 C; 170 G; 148 T; 1 other;

Query Match 0.8%; Score 22; DB 24; Length 711;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGT 78
|||||
DB 48 GTGACCCACGCGTCCGGAGT 69

RESULT 74
ABS77128

ID ABS77128 standard; cDNA; 748 BP.

AC ABS77128;

DT 12-DEC-2002 (first entry)

DE Frog embryonic gene sequence Q9925519.

KW Frog; ss; embryonic development; developmental disorder;

KW microarray; cell differentiation.

OS Xenopus laevis.

PN US2002081610-A1.

PD 27-JUN-2002.

PF 23-JUL-2001; 2001US-0910943.

PR 21-JUL-2000; 2000US-219658P.

PA (UYRQ) UNIV ROCKEFELLER.

PI Hemmati-Briavanlou A, Altman CR;

DR WPI; 2002-626534/67.

PT Nucleic acid array containing Xenopus embryonic nucleic acids is useful
PT to identify genes involved in embryonic development, to identify
PT different types of embryonic cells, and to diagnose developmental
PT disorders -

PS Claim 1; Page 398; 823pp; English.

XX The invention relates to a nucleic acid array, where each coordinate
XX contains a single nucleic acid species having one of 770 nucleotide
XX sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene
XX product, or its complement or hybridisable fragment of not less than
XX 20 contiguous nucleotides of one of those sequences. Also included are
XX detecting differential expression of embryonic genes, comprising:

CC (a) contacting a nucleic acid array comprising genes expressed in
CC embryonic but not mature cells with nucleic acids from sample and
CC control cells; and (b) detecting differential hybridisation of nucleic
CC acids from the sample cells relative to the control cells; and detecting
CC defects in development, comprising: (a) contacting nucleic acids from
CC test cells undergoing development with a nucleic acid array of gene
CC products known to play a fundamental role in the development process; and
CC (b) detecting a difference in expression of a fundamental gene in the
CC sample cells relative to a standard. The invention is useful to identify
CC genes involved in embryonic development and related processes such as
CC cell differentiation. This would be useful for diagnosing developmental
CC disorders and for identifying different types of embryonic cells.
CC The present sequence is one of the 770 Xenopus embryonic cDNA sequences.
XX

SO Sequence 748 BP; 237 A; 122 C; 137 G; 216 T; 36 other;

Query Match 0.8%; Score 22; DB 24; Length 748;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGT 78
|||||
DB 49 GTGACCCACGCGTCCGGAGT 70

RESULT 75
AAA87682

ID AAA87682 standard; cDNA; 2112 BP.

AC AAA87682;

DT 04-DEC-2000 (first entry)

DE Human secreted protein gene 17 SEQ ID NO:27.

KW Human; secreted protein; immunosuppressive; immunostimulant; nootropic;

KW antiinflammatory; cardiant; vulnerary; antileuc; anticonvulsant;

KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytoskeletal;

KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;

KW gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;

KW cancer; immune system disorder; hyperproliferative disorder; infection;

KW cardiovascular disorder; neurological disease; wound healing; ss.

OS Homo sapiens.

PN WO200043495-A2.

PD 27-JUL-2000.

PF 18-JAN-2000; 2000WO-US00903.

PR 19-JAN-1999; 99US-0116330.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;

DR Komatsu G, Birse CB;

DR WPI; 2000-499225/44.

DR P-PSDB; AAB25681.

PT New isolated polynucleotide encoding a secreted protein useful for
PT preventing, treating or ameliorating a medical condition -

PS Claim 1; Page 388; 451pp; English.

XX The polynucleotide sequences given in AAA87666 to AAA87708 encodes the
XX human secreted proteins given in AAB25665 to AAB25755. Human secreted
XX proteins have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: immunosuppressive;
XX immunostimulant; antiinflammatory; cardiant; vulnerary; antileuc;
XX nootropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
XX antibacterial; antiparasitic; thrombolytic; anticoagulant;

CC antiarteriosclerotic and cytostatic. The secreted proteins and their
CC polynucleotides can be used in gene therapy and as vaccines,
CC chemotaxis-modulators and angiogenesis- modulators. The human secreted
CC proteins and polynucleotides can be used for diagnosing (the
CC susceptibility to) a pathological condition by determining the presence
CC or absence of a mutation in the polynucleotide or determining the
CC presence or amount of expression of the protein. The polynucleotides and
CC proteins can also be used in the treatment and diagnosis of cancer,
CC diseases of the immune system, hyperproliferative disorders,
CC cardiovascular disorders and neurological disease. They can also be used
CC to promote wound healing and to fight infection. AA87657 to AA87665 and
CC AA85664 represent sequences used in the exemplification of the present
CC invention.
XX

SQ Sequence 2112 BP; 512 A; 450 C; 480 G; 657 T; 13 other;

Query Match 0.8%; Score 22; DB 21; Length 2112;
Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGT 78
|||
DB 71 GTGACCCACGCGTCCGGAGT 92

Search completed: July 31, 2003, 19:48:34
Job time : 685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 22:28:41 ; Search time 596 Seconds
(without alignments)
9376.961 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aactcttattatcagatg.....aggctttttcctaacc 2709

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1439767 seqs, 1031500376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database: Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCOT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCOT_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	US-09-774-490-1	Sequence 1, Appli
2	2508	92.6	2530	US-10-097-340-283	Sequence 283, App
3	2508	92.6	2530	US-10-262-538-9	Sequence 9, Appli
4	405	15.0	456	US-09-864-761-30992	Sequence 30992, A
5	303	11.2	354	US-09-922-217-736	Sequence 736, App
6	303	11.2	354	US-09-833-263-736	Sequence 736, App
7	292	10.8	354	US-10-025-380-736	Sequence 736, App
8	292	10.8	496	US-09-864-761-14444	Sequence 14444, A
9	160	5.9	172	US-09-864-761-27943	Sequence 27943, A
10	160	5.9	484	US-09-864-761-11363	Sequence 11363, A
11	144	5.3	446	US-09-864-761-11810	Sequence 11810, A
12	123	4.5	123	US-09-864-761-28381	Sequence 28381, A
13	122	4.5	456	US-09-864-761-6056	Sequence 6056, App
14	101	3.7	101	US-09-864-761-22893	Sequence 22893, A
15	71	2.6	2137	US-09-774-490-6	Sequence 6, Appli
16	67	2.5	520	US-10-027-632-238542	Sequence 238542,

17	67	2.5	520	13	US-10-027-632-238543	Sequence 238543,
18	31	1.1	4661	10	US-09-897-006-10	Sequence 10, Appli
19	31	1.1	4661	11	US-09-897-511A-10	Sequence 10, Appli
20	31	1.1	7160	8	US-08-786-531B-5	Sequence 5, Appli
21	31	1.1	7235	8	US-08-786-531B-6	Sequence 6, Appli
22	26	1.0	1313	10	US-09-728-422-1	Sequence 1, Appli
23	25	0.9	1400	9	US-09-774-490-5	Sequence 5, Appli
24	22	0.8	115	14	US-10-106-698-3299	Sequence 3299, App
25	22	0.8	711	9	US-09-910-943-236	Sequence 236, App
26	22	0.8	748	9	US-09-910-943-361	Sequence 361, App
27	22	0.8	5732	10	US-09-897-006-6	Sequence 6, Appli
28	22	0.8	5732	11	US-09-897-511A-6	Sequence 6, Appli
29	22	0.8	7352	8	US-08-786-531B-4	Sequence 4, Appli
30	22	0.8	7353	10	US-08-786-531B-1	Sequence 1, Appli
31	22	0.8	9183	8	US-09-897-006-7	Sequence 7, Appli
32	22	0.8	9183	11	US-09-897-511A-7	Sequence 7, Appli
33	21	0.8	332	10	US-09-924-035A-528	Sequence 528, App
34	21	0.8	442	10	US-09-925-300-828	Sequence 828, App
35	21	0.8	457	10	US-09-960-352-4373	Sequence 4373, App
36	21	0.8	690	13	US-10-027-632-235696	Sequence 235696,
37	21	0.8	717	14	US-10-106-698-1856	Sequence 1858, App
38	21	0.8	725	9	US-09-910-943-67	Sequence 67, Appli
39	21	0.8	730	9	US-09-910-943-448	Sequence 448, App
40	21	0.8	745	9	US-09-910-943-674	Sequence 674, App
41	21	0.8	749	9	US-09-910-943-359	Sequence 359, App
42	21	0.8	749	9	US-09-910-943-428	Sequence 428, App
43	21	0.8	750	9	US-09-910-943-365	Sequence 365, App
44	21	0.8	750	9	US-09-910-943-364	Sequence 504, App
45	21	0.8	781	9	US-09-910-943-383	Sequence 383, App
46	21	0.8	784	9	US-09-910-943-609	Sequence 609, App
47	21	0.8	883	14	US-10-198-846-14063	Sequence 1063, A
48	21	0.8	1181	14	US-10-037-270-33	Sequence 33, Appli
49	21	0.8	1280	14	US-10-106-698-1342	Sequence 1342, App
50	21	0.8	1385	9	US-09-925-776-1	Sequence 1, Appli
51	21	0.8	1394	14	US-10-119-926-44	Sequence 44, Appli
52	21	0.8	1572	14	US-10-037-270-664	Sequence 664, App
53	21	0.8	1808	9	US-09-838-573-1	Sequence 1, Appli
54	21	0.8	1808	14	US-10-160-501-1	Sequence 1, Appli
55	21	0.8	2570	9	US-09-925-302-192	Sequence 192, App
56	21	0.8	2810	11	US-09-786-753-89	Sequence 89, Appli
57	21	0.8	2815	9	US-09-782-960-67	Sequence 67, Appli
58	21	0.8	4207	10	US-09-897-006-4	Sequence 4, Appli
59	21	0.8	4207	11	US-09-897-511A-4	Sequence 4, Appli
60	21	0.8	4210	10	US-09-897-006-5	Sequence 5, Appli
61	21	0.8	4210	11	US-09-897-511A-5	Sequence 5, Appli
62	21	0.8	6255	10	US-09-897-006-13	Sequence 13, Appli
63	21	0.8	6255	11	US-09-897-511A-13	Sequence 13, Appli
64	21	0.8	9511	10	US-09-897-006-34	Sequence 34, Appli
65	21	0.8	9511	11	US-09-897-511A-34	Sequence 34, Appli
66	20	0.7	100	14	US-10-106-698-4090	Sequence 4090, App
67	20	0.7	122	11	US-09-764-891-522	Sequence 522, App
68	20	0.7	142	14	US-10-106-698-4018	Sequence 4018, App
69	20	0.7	197	14	US-10-106-698-2365	Sequence 2365, App
70	20	0.7	327	11	US-09-918-995-8601	Sequence 8601, App
71	20	0.7	355	11	US-09-986-480-113	Sequence 113, App
72	20	0.7	367	14	US-10-198-846-103302	Sequence 10302, A
73	20	0.7	402	10	US-09-925-300-802	Sequence 802, App
74	20	0.7	405	14	US-10-106-698-21772	Sequence 2172, App
75	20	0.7	430	10	US-09-960-352-13318	Sequence 13318, A
76	20	0.7	435	9	US-09-764-869-326	Sequence 326, App
77	20	0.7	435	14	US-10-091-504-326	Sequence 326, App
78	20	0.7	445	14	US-10-106-698-229	Sequence 229, App
79	20	0.7	453	11	US-09-918-995-7946	Sequence 7946, App
80	20	0.7	464	11	US-09-918-995-8118	Sequence 8118, App
81	20	0.7	474	14	US-10-102-806-41	Sequence 41, Appli
82	20	0.7	475	11	US-09-918-995-6137	Sequence 6137, App
83	20	0.7	481	11	US-09-918-995-7696	Sequence 7696, App
84	20	0.7	489	9	US-09-908-711-49	Sequence 49, Appli
85	20	0.7	481	11	US-09-764-891-567	Sequence 567, App
86	20	0.7	486	9	US-09-925-301-232	Sequence 232, App
87	20	0.7	491	11	US-09-918-995-7344	Sequence 7344, App
88	20	0.7	533	10	US-09-924-035A-540	Sequence 540, App
89	20	0.7	556	9	US-09-908-711-51	Sequence 51, Appli

90	20	0.7	556	11	US-09-986-480-64	Sequence 670, App
91	20	0.7	556	11	US-09-764-891-570	Sequence 570, App
92	20	0.7	557	14	US-10-102-806-198	Sequence 198, App
93	20	0.7	558	13	US-10-027-652-13709	Sequence 13709, App
94	20	0.7	606	10	US-09-764-877-115	Sequence 115, App
95	20	0.7	612	11	US-09-986-480-55	Sequence 55, App
96	20	0.7	692	10	US-09-764-984-429	Sequence 429, App
97	20	0.7	693	10	US-10-106-658-1769	Sequence 1769, App
98	20	0.7	696	9	US-09-910-943-288	Sequence 288, App
99	20	0.7	702	9	US-09-910-943-240	Sequence 240, App
100	20	0.7	717	9	US-09-910-943-123	Sequence 120, App

ALIGNMENTS

```

RESULT 1
US-09-774-490-1
; Sequence 1, Application US/09774490
; Patent No. US2001003432A1
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 0733-118001
; CURRENT APPLICATION NUMBER: US/09/774,490
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1

```

Query Match	100.0%;	Score 2709;	DB 9;	Length 2709;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2709; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Oy	AAATCTTTATTTTATGCAATGTTAAACAGCTGTGTATGCAATCCACAGCTGAGGGGTGCG	60
Db	1 AATCTTTATTTTATGCAATGTTAAACAGCTGTGTATGCAATCCACAGCTGAGGGGTGCG	60
Oy	61 ACCACGCGTCCGGAGTAGTGAAGCTCGCTGCTTCCATTGTCCAGCACTAATT	120
Db	61 ACCACGCGTCCGGAGTAGTGAAGCTCGCTGCTTCCATTGTCCAGCACTAATT	120
Oy	121 CGAGATTGTTGAACCTCTCTGGCGGCAAAATACAGGAAGAAAGACTAAACAGCAAAAG	180
Db	121 CGAGATTGTTGAACCTCTCTGGCGGCAAAATACAGGAAGAAAGACTAAACAGCAAAAG	180
Oy	181 GACCTACAGCGCTGCGAGCATGGGCTGTGTTAACTAGATTGTCTCTCTTTCTGCGGAGT	240
Db	181 GACCTACAGCGCTGCGAGCATGGGCTGTGTTAACTAGATTGTCTCTCTTTCTGCGGAGT	240
Oy	241 ATTAATCTTACAGCAAGCAAACTATAGAAATGGGAAGAAACAATGTGCCAAGCTGAATT	300
Db	241 ATTAATCTTACAGCAAGCAAACTATAGAAATGGGAAGAAACAATGTGCCAAGCTGAATT	300
Oy	301 ATCTCTCAAAAGAAATTTGGAATCCAAACAATGTGATCACTTTCATGTGCTGGCCCAACAG	360
Db	301 ATCTCTCAAAAGAAATTTGGAATCCAAACAATGTGATCACTTTCATGTGCTGGCCCAACAG	360
Oy	361 CTCGAGTTATCATACCTCTCTTTGGATGAGGAACGAGTAGAGCTGTATGTTGGACAAA	420
Db	361 CTCGAGTTATCATACCTCTCTTTGGATGAGGAACGAGTAGAGCTGTATGTTGGACAAA	420
Oy	421 GGAATCAATTTTTTCATGTGCACTGGTTAATTCAGGAATTTTCAAAAGATTTGTGTGCC	480
Db	421 GGAATCAATTTTTTCATGTGCACTGGTTAATTCAGGAATTTTCAAAAGATTTGTGTGCC	480
Oy	481 AGTATCTTACACCAAGAAAGATGAATGCAAGTGGGCTGGAAAAGACATCTGAAAGAATG	540

Db	4481	AGTAATCTTACACGAGAAAGATGAAATGCAATGGGCTGAAAAAGCATCTGAAAGAAATG	540
Qy	541	TGCTAATTTCAATCAAGGTAATTAAGGCAATATATACATCTGTAAGCCTGTGGAAC	600
Db	541	TGCTAATTTCAATCAAGGTAATTAAGGCAATATATACATCTGTAAGCCTGTGGAAC	600
Qy	601	GGGGGCTTTTCAATCCAAATTTGCAACCTAATATGAAATTTGCAATCACTCTGAGCAAAAT	660
Db	601	GGGGGCTTTTCAATCCAAATTTGCAACCTAATATGAAATTTGCAATCACTCTGAGCAAAAT	660
Qy	661	TTTTAAGCTGAGAACTCACATTTTGAAGAGCGCGTGGAGAGAGTCATATGACCTTAA	720
Db	661	TTTTAAGCTGAGAACTCACATTTTGAAGAGCGCGTGGAGAGAGTCATATGACCTTAA	720
Qy	721	GCTGCTGACAGACATCCCTTTTAATAGATGAGAAATTAATATCTGTGAACTGACGTAAT	780
Db	721	GCTGCTGACAGACATCCCTTTTAATAGATGAGAAATTAATATCTGTGAACTGACGTAAT	780
Qy	781	TATGGGGGGAACATTTTGCTATCTTCCGAACTCTTGGGGCAACCAACCCAAATCAGACGA	840
Db	781	TATGGGGGGAACATTTTGCTATCTTCCGAACTCTTGGGGCAACCAACCCAAATCAGACGA	840
Qy	841	GCAGATATTTCCAGGTGGCTCAATGATCCAAAGTTCAATPAAGTGGCCACTCATCTCAGA	900
Db	841	GCAGATATTTCCAGGTGGCTCAATGATCCAAAGTTCAATPAAGTGGCCACTCATCTCAGA	900
Qy	901	GAGTGACAAATCCTGAGAGATGACAAAGTATATCTTTTCTTCCGTGAAAATGCAATAGATG	960
Db	901	GAGTGACAAATCCTGAGAGATGACAAAGTATATCTTTTCTTCCGTGAAAATGCAATAGATG	960
Qy	961	AGAACACTCTGTGAAAAGCTATCTACGCTGAAATATGCTCAGATATGCAAGAAATGACTTTGG	1020
Db	961	AGAACACTCTGTGAAAAGCTATCTACGCTGAAATATGCTCAGATATGCAAGAAATGACTTTGG	1020
Qy	1021	AGGGCAGCAAGAGTCTGTGAATTAATGAGCAAACTCCTCAAGCTGCTGATTTGGCTC	1080
Db	1021	AGGGCAGCAAGAGTCTGTGAATTAATGAGCAAACTCCTCAAGCTGCTGATTTGGCTC	1080
Qy	1081	AGTGCAGAGTCCAAATGAGCATTTGATGAACTGAGATGCAAGTATGATTTCTAAT	1140
Db	1081	AGTGCAGAGTCCAAATGAGCATTTGATGAACTGAGATGCAAGTATGATTTCTAAT	1140
Qy	1141	GAACTTTAAGATCTCTAATAATTCAGATTGATATGAGAGTGTTAACATTCACAGTAACAT	1200
Db	1141	GAACTTTAAGATCTCTAATAATTCAGATTGATATGAGAGTGTTAACATTCACAGTAACAT	1200
Qy	1201	TTTCAAGGGAATCAACCCGTGTATGTATGATAGATAGATGATGAGAAAGGAGTTCCTTGG	1260
Db	1201	TTTCAAGGGAATCAACCCGTGTATGTATGATAGATAGATGATGAGAAAGGAGTTCCTTGG	1260
Qy	1261	TCCATATATGCCACAGGATGAGACCCAACTATCAATGAGGTGCTTATCAGAGAAAGTCC	1320
Db	1261	TCCATATATGCCACAGGATGAGACCCAACTATCAATGAGGTGCTTATCAGAGAAAGTCC	1320
Qy	1321	CTATTCACAGGCGCAAGAACTTGTCACAGCAAAACATTTGGTGTTTGACTCTACAAAGGA	1380
Db	1321	CTATTCACAGGCGCAAGAACTTGTCACAGCAAAACATTTGGTGTTTGACTCTACAAAGGA	1380
Qy	1381	CCTTCCTGATGATGTTATTAACCTTTGCAAGAGTCATCCAGCCATGTCATCCAGTGT	1440
Db	1381	CCTTCCTGATGATGTTATTAACCTTTGCAAGAGTCATCCAGCCATGTCATCCAGTGT	1440
Qy	1441	TCCTATGAAACAATGGCCCAATAGATCAAAAGGAGTAAATTAATCAATTTACAAAT	1500
Db	1441	TCCTATGAAACAATGGCCCAATAGATCAAAAGGAGTAAATTAATCAATTTACAAAT	1500
Qy	1501	TGTGCTAGACCGAGTGTGATGCAAGAGATGACAGTATGATGTTATGTTATCGGAACGA	1560
Db	1501	TGTGCTAGACCGAGTGTGATGCAAGAGATGACAGTATGATGTTATGTTATCGGAACGA	1560
Qy	1561	TGTTGGAGCGTCTTAAAGTATTTCAAATTCCTAAGAGACTTGGTATGATTTAAGAGA	1620
Db	1561	TGTTGGAGCGTCTTAAAGTATTTCAAATTCCTAAGAGACTTGGTATGATTTAAGAGA	1620


```

Db      1561 TGTGGAGCCGTTCTTAAAGTAGTTCAATTCCTTAAGAGACCTTGATGATTGAAGA 1620
Qy      1621 GGTTCGCTGGAAGAAATGACAGTTTTTCGGGACCGACTGCTATTTGACAAATGAGACT 1680
Db      1621 GGTTCGCTGGAAGAAATGACAGTTTTTCGGGACCGACTGCTATTTGACAAATGAGACT 1680
Qy      1681 TTCCACTAAGCAGCAACATATATATGTTCAACGGCTGGGGTTGCCAGCTCCCTTT 1740
Db      1681 TTCCACTAAGCAGCAACATATATATGTTCAACGGCTGGGGTTGCCAGCTCCCTTT 1740
Qy      1741 ACACCGGCTGATATTTACGGGAAACGCTGCTGATGTTGCTCGCCCGGAGACCTTA 1800
Db      1741 ACACCGGCTGATATTTACGGGAAACGCTGCTGATGTTGCTCGCCCGGAGACCTTA 1800
Qy      1801 CTGTCCTGGGATGTTCTGCAATGTTCTCGATTTTCCACTGCAAGAGACGACAAAG 1860
Db      1801 CTGTCCTGGGATGTTCTGCAATGTTCTCGATTTTCCACTGCAAGAGACGACAAAG 1860
Qy      1861 ACGACAGATATTAAGAAATGAGACCCCACTGACTGTTTCAAGCTTACACATGATTA 1920
Db      1861 ACGACAGATATTAAGAAATGAGACCCCACTGACTGTTTCAAGCTTACACATGATTA 1920
Qy      1921 TCACATGCGCCACAGCCCTGAAAGAGATATCTATGCTGTAAGAAATGATGACACTT 1980
Db      1921 TCACATGCGCCACAGCCCTGAAAGAGATATCTATGCTGTAAGAAATGATGACACTT 1980
Qy      1981 TTAGGATGAGAGTCCGAGAGTCCGACAGAGCGCTGCTATTGGGCAATTCCAGGCGGAA 2040
Db      1981 TTAGGATGAGAGTCCGAGAGTCCGACAGAGCGCTGCTATTGGGCAATTCCAGGCGGAA 2040
Qy      2041 TGAAGAGCGAAAGAGAGATCAGATGATGATCATATCATCAGACAGATCAAGGCT 2100
Db      2041 TGAAGAGCGAAAGAGAGATCAGATGATGATCATATCATCAGACAGATCAAGGCT 2100
Qy      2101 TCTGCTAGTGTCTACACAGAGAGATTCAGGCAATTAACCTCTGCTCGGCTGGAACA 2160
Db      2101 TCTGCTAGTGTCTACACAGAGAGATTCAGGCAATTAACCTCTGCTCGGCTGGAACA 2160
Qy      2161 TGGGTTCAATCAAACTTCTTAAGTAACTCCGAGATCAATGACAGAGCAATTTGGA 2220
Db      2161 TGGGTTCAATCAAACTTCTTAAGTAACTCCGAGATCAATGACAGAGCAATTTGGA 2220
Qy      2221 AGAATCTTCTAATAAGATGATGATGATGATGCTCTAAGACCAAGAAATGCTCAATG 2280
Db      2221 AGAATCTTCTAATAAGATGATGATGATGATGCTCTAAGACCAAGAAATGCTCAATG 2280
Qy      2281 CATGACCTTACCGCAGAAAGTCTGCTACAGAGACTTCAATGACCTCAACACCCCA 2340
Db      2281 CATGACCTTACCGCAGAAAGTCTGCTACAGAGACTTCAATGACCTCAACACCCCA 2340
Qy      2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db      2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Qy      2401 GCAAGAGCAGAGACATACCCAGGAGACGTAACAAATGAGCACTTACAAAGAAATTA 2460
Db      2401 GCAAGAGCAGAGACATACCCAGGAGACGTAACAAATGAGCACTTACAAAGAAATTA 2460
Qy      2461 GAAAGGTGAAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTCTGAGCTGC 2520
Db      2461 GAAAGGTGAAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTCTGAGCTGC 2520
Qy      2521 ATTACCTTAGAACTCAAACTAGAACTTCCCTAGACAAATTAACCTGGAAGAAACAA 2580
Db      2521 ATTACCTTAGAACTCAAACTAGAACTTCCCTAGACAAATTAACCTGGAAGAAACAA 2580
Qy      2581 TGCATATATCAATGAACTTTTTCATGCAATTAATGATGATGATGATGATGATGATGAT 2640
Db      2581 TGCATATATCAATGAACTTTTTCATGCAATTAATGATGATGATGATGATGATGATGAT 2640
Qy      2641 AGCTGAGTTCCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
Db      2641 AGCTGAGTTCCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700

```

```

Qy      2701 CCTAATACC 2709
Db      2701 CCTAATACC 2709

RESULT 2
US-10-097-340-283
; Sequence 283, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANVARAPU
; APPLICANT: Sebastian HOERSCHE
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT FILING DATE: 2002-03-14
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-283

Query Match      92.6%; Score 2508; DB 14; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY		433	TTTCATTCCGACCTGGTAAATATCAAGATTTTCAAAGAATGTGTGCCAGTACTCTTAAC	492
Db		249	TTTCATTCCGACCTGGTAAATATCAAGATTTTCAAAGAATGTGTGCCAGTACTCTTAAC	308
OY		493	CAGAAGATGGAATGCMACTGGCGCTGGAAAAGACA TCTGAAAGAATGTGCTAATTTTCAT	552
Db		309	CAGAAGATGGAATGCMACTGGCGCTGGAAAAGACA TCTGAAAGAATGTGCTAATTTTCAT	368
OY		553	CAAGGTACTTAAGGCATATATCAGACTCACCTTGTAACGCTGTGGAAAGGGGGCTTTTCA	612
Db		369	CAAGGTACTTAAGGCATATATCAGACTCACCTTGTAACGCTGTGGAAAGGGGGCTTTTCA	428
OY		613	TCCAATTTGCACTACATTTGAATAATTTGACATCATCTGAGAGCAATATTTTTAAGCTGA	672
Db		429	TCCAATTTGCACTACATTTGAATAATTTGACATCATCTGAGAGCAATATTTTTAAGCTGA	488
OY		673	GAACTCACATTTTGA AAAACGGCGCTGGGAAGATCCATATGACCCTTAAGCTGTGCACGC	732
Db		489	GAACTCACATTTTGA AAAACGGCGCTGGGAAGATCCATATGACCCTTAAGCTGTGCACGC	548
OY		733	ATCCCTTTTAAATAGATGGAATTTATCTGTGGAACTGACGTGATTTTATGGGGCCAGA	792
Db		549	ATCCCTTTTAAATAGATGGAATTTATCTGTGGAACTGACGTGATTTTATGGGGCCAGA	608
OY		793	CTTTGCTATCTTCCGCAACTCTTGGGCA CCA CACC CAATCAGGA CAGAGCAGATATTC	852
Db		609	CTTTGCTATCTTCCGCAACTCTTGGGCA CCA CACC CAATCAGGA CAGAGCAGATATTC	668
OY		853	CAGGTGGCTCAATGATCCAAAGTTCATTATGTGCCCACTCATCTCAGAGATGCATAATCC	912
Db		669	CAGGTGGCTCAATGATCCAAAGTTCATTATGTGCCCACTCATCTCAGAGATGCATAATCC	728
OY		913	TGAATATGACAAAGATATCTTTTCTCCGGA AAATGCATATGATGAGAAACCTCGG	972
Db		729	TGAATATGACAAAGATATCTTTTCTCCGGA AAATGCATATGATGAGAAACCTCGG	788
OY		973	AAAACTACTCAGCTAGATAGCTCAGATATGCAAGATGACTTTGGAGGGCACAGANG	1032
Db		789	AAAACTACTCAGCTAGATAGCTCAGATATGCAAGATGACTTTGGAGGGCACAGANG	848
OY		1033	TCTGTGTAATTAATGAGACA CATCTCTCAAAAGCTGCTGATTTGCTCAGTGCCAGGTC	1092
Db		849	TCTGTGTAATTAATGAGACA CATCTCTCAAAAGCTGCTGATTTGCTCAGTGCCAGGTC	908
OY		1093	AAATGGCACTGACATCTCAATTTTGATGA ACTGAGGATGTATTTCCAAAGAACCTTAAAGA	1152
Db		909	AAATGGCACTGACATCTCAATTTTGATGA ACTGAGGATGTATTTCCAAAGAACCTTAAAGA	968
OY		1153	TCTTAAAAATCCAGTTGTATATGAGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC	1212
Db		969	TCTTAAAAATCCAGTTGTATATGAGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC	1028
OY		1213	AGCGCTGTGTATGTATAGACATGAGTGAAGTGAAGGGGTGTTCTTGGTCAATATGCCCA	1272
Db		1029	AGCGCTGTGTATGTATAGACATGAGTGAAGTGAAGGGGTGTTCTTGGTCAATATGCCCA	1088
OY		1273	CAGGGATGAGCCCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCCTATCCACGGCC	1332
Db		1089	CAGGGATGAGCCCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCCTATCCACGGCC	1148
OY		1333	AGGAATTTGTCCACAGCAAAACATTTGTGTGTTTGACACTTCAAAAGACCTTCTGATGA	1392
Db		1149	AGGAATTTGTCCACAGCAAAACATTTGTGTGTTTGACACTTCAAAAGACCTTCTGATGA	1208
OY		1393	TGTTATTAACCTTTTGCAGAAAGTCA TCAGCCATGTACAATCAATCAGTGTTCCTATGAACA	1452
Db		1209	TGTTATTAACCTTTTGCAGAAAGTCA TCAGCCATGTACAATCAATCAGTGTTCCTATGAACA	1268
OY		1453	TGCGCCAAATAGATCAAAACGGAATGTAAATTAATCAATTTACAAAATTTGTGTGAGACCG	1512
Db		1269	TGCGCCAAATAGATCAAAACGGAATGTAAATTAATCAATTTACAAAATTTGTGTGAGACCG	1328

QY	1513	AGTGGATCCAGAAAGTGGACAGTATGATGTTATCGAAACAGATGTGGACCGT	1572
Db	1329	AGTGATCCAGAAATGGACAGTATGATGTTATCGAAACAGATGTGGACCGT	1388
QY	1573	TCTTAAAGTAGTTCAATTCTTAAGAGACTTGGTATGATTTAGAAAGGTTCTGTGGGA	1632
Db	1389	TCTTAAAGTAGTTCAATTCTTAAGAGACTTGGTATGATTTAGAAAGGTTCTGTGGGA	1448
QY	1633	AGAAATGACAGTTTTTGGGAAACCGACTGCTATTTTCAGCATTTGAGAGCTTTCACATMAACA	1692
Db	1449	AGAAATGACAGTTTTTGGGAAACCGACTGCTATTTTCAGCAATTTGAGAGCTTTCACATMAACA	1508
QY	1693	GCAACCAATATATATATGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGGA	1752
Db	1509	GCAACCAATATATATATGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGGA	1566
QY	1753	TATTTACGGGAAAGCGTGTGCTAGTGTGCTCGCCCGAGACCCCTTAAGTGTCTGGGA	1812
Db	1569	TATTTACGGGAAAGCGTGTGCTAGTGTGCTCGCCCGAGACCCCTTAAGTGTCTGGGA	1628
QY	1813	TGTTTCTGCAAGTTCTGTGCTATTTTCCCACTGCAAGAGACGCACAGACGACAAGATAT	1872
Db	1629	TGTTTCTGCAAGTTCTGTGCTATTTTCCCACTGCAAGAGACGCACAGACGACAAGATAT	1688
QY	1873	AAGAAATGGAAACCCACTGACTCATCTGTTCAGACTTACACCATATATATCAACATGGCCA	1932
Db	1689	AAGAAATGGAAACCCACTGACTCATCTGTTCAGACTTACACCATATATATCAACATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATCTATGTGTAGAAATAGTAGCAATTTTGGAAATGCAG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTATGTGTAGAAATAGTAGCAATTTTGGAAATGCAG	1808
QY	1993	TCCGAAGTCGACAGAGAGCGCTGTGCTATTTGGCAATTCACAGCGGAAATGAAGCGAAA	2052
Db	1809	TCCGAAGTCGACAGAGAGCGCTGTGCTATTTGGCAATTCACAGCGGAAATGAAGCGAAA	1868
QY	2053	AGAAAGATCAGAGAGGATGATTCATATCATCAGACAGATCAAGAGCCCTTCGTAACGAG	2112
Db	1869	AGAAAGATCAGAGAGGATGATTCATATCATCAGACAGATCAAGAGCCCTTCGTAACGAG	1928
QY	2113	TCTAACACAGAGAGATTCAGGCAATTAACCTTCGCATGCGGTGGAAACATGGGTTCATACA	2172
Db	1929	TCTAACACAGAGAGATTCAGGCAATTAACCTTCGCATGCGGTGGAAACATGGGTTCATACA	1988
QY	2173	AACTCTTTTAAAGTAAACCTGTGAAGTCATTGACACAGAGCATTTGGAAAGACTTCTTCA	2232
Db	1989	AACTCTTTTAAAGTAAACCTGTGAAGTCATTGACACAGAGCATTTGGAAAGACTTCTTCA	2048
QY	2233	TAAAGATGATGATGAGATGGCTTAAGACCAAGAAATGTCCAAATGACATGACACCTAG	2292
Db	2049	TAAAGATGATGATGAGATGGCTTAAGACCAAGAAATGTCCAAATGACATGACACCTAG	2108
QY	2293	CCAGAAGGTCGTATACAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAAGGTCGTATACAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2168
QY	2353	GGATGAGTTCTGTGAACAAGTTTGGAAAGGGAACGAAACACGTCGGCAAAGCCAGG	2412
Db	2169	GGATGAGTTCTGTGAACAAGTTTGGAAAGGGAACGAAACACGTCGGCAAAGCCAGG	2228
QY	2413	ACATATCCCGAGGGAACAGTAAACAAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAA	2472
Db	2229	ACATATCCCGAGGGAACAGTAAACAAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAA	2288
QY	2473	CAGAGAGACCCAGAAATTTTGAAGAGGGAACCCAGAGAGTCTGAGCTGCATTACTTGA	2532
Db	2289	CAGAGAGACCCAGAAATTTTGAAGAGGGAACCCAGAGAGTCTGAGCTGCATTACTTGA	2348
QY	2533	AACTCAAAACAATGAGAACTTGTCTAGACAAATTAAGTGAAGAAACAAATGCAATATACAT	2592
Db	2349	AACTCAAAACAATGAGAACTTGTCTAGACAAATTAAGTGAAGAAACAAATGCAATATACAT	2408
QY	2593	GAACTTTTTCATGCAATTATGTGATGTTTACATGCTGGGAAATTCAGCTGAGTTCCA	2652

1689 AAGAAATGAGAGACCCCTGACTGCTTCAAGACTTACACATGATATACCATGAGCA 1748
1933 CAGCCCTGAAGAGAGATATCTATGTGTAGAGAAATAGATACATTTTGAATGAG 1992
1749 CAGCCCTGAAGAGAGATATCTATGTGTAGAGAAATAGATACATTTTGAATGAG 1808
1993 TCCGAAGTGCAGAGAGAGCGCTGTCTATTGGCAATTCAGAGGCGAAATGAGAGCA 2052
1809 TCCGAAGTGCAGAGAGAGCGCTGTCTATTGGCAATTCAGAGGCGAAATGAGAGCA 1868
2053 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2112
1869 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1928
2113 TCTACAAAGAAAGATTCAGAGGCAATTCCTGCTCCATGCGGTGGAAATGAGGTCTAT 2172
1929 TCTACAAAGAAAGATTCAGAGGCAATTCCTGCTCCATGCGGTGGAAATGAGGTCTAT 1988
2173 AACTCTTCTTAAGTAACTCTGAAATGATGATGATGATGATGATGATGATGATGAT 2232
1989 AACTCTTCTTAAGTAACTCTGAAATGATGATGATGATGATGATGATGATGATGAT 2048
2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2292
2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2108
2293 CGAAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
2109 CGAAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2168
2353 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2412
2169 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228
2413 ACATATCCCAAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2472
2229 ACATATCCCAAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2288
2473 CAGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2532
2289 CAGGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2348
2533 AACCTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2592
2349 AACCTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2408
2593 GAACCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
2409 GAACCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
2653 CCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
2469 CCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2516

RESULT 4
US-09-864-761-30992
Sequence 30992, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomicca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30992
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
OTHER INFORMATION: EST HUMAN HIT: BF700780.1, EVALU 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 2.00e-88
OTHER INFORMATION: NT HIT: g11421514, EVALU 0.00e+00
US-09-864-761-30992

Query Match 15.0%; Score 405; DB 9; Length 456;
Best Local Similarity 99.8%; Pred. No. 1.2e-211;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2060 ATCAGAGTGAATGATCATATATATCAGAGACAGATCAAGGCTTCTGCTACGATGCTACA 2119
1 ATCAGAGTGAATGATCATATATATCAGAGACAGATCAAGGCTTCTGCTACGATGCTACA 60
2120 CAGAAGATTCAGGCAATTAATCCTGCGCATCGGTGGAACATGGGTTCATACAACTCTT 2179
61 CAGAAGATTCAGGCAATTAATCCTGCGCATCGGTGGAACATGGGTTCATACAACTCTT 120
2180 CTTAAGGTAACTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTCATTAAGAT 2239
121 CTTAAGGTAACTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTCATTAAGAT 180
2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATGACATGACACCTTACGAGAG 2299
181 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATGACATGACACCTTACGAGAG 240
2300 GTCTGGTACAGAGCTTCATGACAGCTCATCAACACCCCAATCTCAACAGATGATGAG 2359
241 GTCTGGTACAGAGCTTCATGACAGCTCATCAACACCCCAATCTCAACAGATGATGAG 300
2360 TTCTGTGAACAAGTTTGGAAAAAGGAGCGAAGAAACAAAGCTGGGCAAGGCCAGACATACC 2419

Db 301 TTCTGTGAACAAGTTGAAAAAGGAGCCGCAAGGCGCAGACATACC 360
Qy 2420 CCAGGAAACAGTAACAATGAGAGCACTTACAGAAAATTAAGAAAGTAAGAAACGAGG 2479
Db 361 CAGGGAACAGTAACAATGAGAGCACTTACAGAAAATTAAGAAAGTAAGAAACGAGG 420
Qy 2480 ACCGAGAAATTTGAGAGGCAACCCAGAGTGTCTGA 2515
Db 421 ACCGAGAAATTTGAGAGGCAACCCAGAGTGTCTGA 456

RESULT 5
US-09-922-217-736
; Sequence 736, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: n = A,T,C or G
US-09-922-217-736

Query Match 11.2%; Score 303; DB 9; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.5e-155;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2098 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 2157
Db 1 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 60
Qy 2158 ACATGGGTTCAATAAACTCTTTTAAGTAAACCTGGAAGTCATTGACAGAGCATTT 2217
Db 61 ACATGGGTTCAATAAACTCTTTTAAGTAAACCTGGAAGTCATTGACAGAGCATTT 120
Qy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCAA 2277
Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCAA 180
Qy 2278 TAGCATGACACTTACGCAAGAGTGTGTACAGAACTTACGAGCTCATCAACCAACC 2337
Db 181 TAGCATGACACTTACGCAAGAGTGTGTACAGAACTTACGAGCTCATCAACCAACC 240
Qy 2338 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 2397
Db 241 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 300
Qy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAACGTAACAATGAGACCTTACA 2451
Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAACGTAACAATGAGACCTTACA 354

RESULT 6
US-09-833-263-736
; Sequence 736, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(354)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-263-736

Query Match 11.2%; Score 303; DB 10; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.5e-155;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2098 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 2157
Db 1 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 60
Qy 2158 ACATGGGTTCAATAAACTCTTTTAAGTAAACCTGGAAGTCATTGACAGAGCATTT 2217
Db 61 ACATGGGTTCAATAAACTCTTTTAAGTAAACCTGGAAGTCATTGACAGAGCATTT 120
Qy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCAA 2277
Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCAA 180
Qy 2278 TAGCATGACACTTACGCAAGAGTGTGTACAGAACTTACGAGCTCATCAACCAACC 2337
Db 181 TAGCATGACACTTACGCAAGAGTGTGTACAGAACTTACGAGCTCATCAACCAACC 240
Qy 2338 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 2397
Db 241 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 300
Qy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAACGTAACAATGAGACCTTACA 2451
Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAACGTAACAATGAGACCTTACA 354

RESULT 7
US-10-025-380-736
; Sequence 736, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun

```

1  APPLICANT: Clapper, Jonathan D.
2  APPLICANT: Skeiky, Yaeli A. W.
3  APPLICANT: Panger, Gary R.
4  APPLICANT: Vedick, Thomas S.
5  APPLICANT: Carter, Darick
6  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
7  TITLE OF INVENTION: OF COLOM CANCER AND METHODS FOR THEIR USE
8  FILE REFERENCE: 210121.471C14
9  CURRENT APPLICATION NUMBER: US/10/025,380
10 CURRENT FILING DATE: 2001-12-19
11 NUMBER OF SEQ ID NOS: 1129
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 736
14     LENGTH: 354
15     TYPE: DNA
16 ORGANISM: Homo sapiens
17 FEATURE:
18 NAME/key: misc_feature
19 LOCATION: 57..
20 OTHER INFORMATION: n = A,T,C or G
21 US-10-025-380-736

```

Query Match	11.2%	Score 303;	DB 13;	Length 354;
Best Local Similarity	99.7%;	Pred No. 1.5e-155;		
Matches 353;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

2038 CCTCTGCGACGTAGTCTACACGAGAAGATTACCTCTGCGCATGGGGTGA 2157
Db 1 CCTCTGCTACGTAGTCTACACGAGAAGATTACCTCTGCGCATGGGGNGA 60

QY 2158 ACATGGGTTTCATCAAACTCTTCTTTAAGTAAACCTCGAAGTCATTGACACAGCATTT 2217
| | | | |
Db 61 ACAATGGTTCATCAAAACTCTTTTAAGGAACCTCGAAGTCATTGACACAGGCATTT 120
| | | | |

Oy 2218 GGAAGAACTTCTTCATAAAGATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAA 22777

b 121 GGAAGAACTTCTTCATAAAGATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAA 180

Oy 2278 TAGCATGACACCTAGCCAGAGGCTGTGTAACAGAGACTTCATGACGCTCATCAACACC 2337
 Db 181 TAGATGACACCTAGCCAGAGGCTGTGTAACAGAGACTTCATGACGCTCATCAACACC 240

OY 2338 CAATCTCAACACGATGATGAGTCTGTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAG 23397
 DB 241 CAATCTCTAAACAGATGGAGTACTTCTGTAAACAAGTTTGGAAAAAGGACCTGAATATCTATCT 300

Oy	2398	T	C	G	G	C	A	A	A	G	G	C	C	A	G	A	C	A	T	A	C	C	C	A	G	G	A	C	A	G	T	A	C	A	A	T	G	A	A	G	A	G	A	C	T	T	A	C	A	
Ob	301	T	C	G	G	C	A	A	A	G	G	C	C	A	G	A	C	A	T	A	C	C	C	A	G	G	A	C	A	G	T	A	C	A	A	T	G	A	A	G	A	G	A	C	T	T	A	C	A	

RESULT 8

US20020048763A1
Sequence 14444, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:

APPLICANT: Ferni, Shaïon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864.761
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USE
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/307,455

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

2000-10-04

```

: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 14444

```

```

? LENGTH: 496
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC006322.2 SIGNAL = 0.73
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
US-09-864-761-1444

```

Query Match	10.8%	Score 292;	DB 9;	Length 496;
Best Local Similarity	100.0%;	Pred. No. 1.8e-149;		
Matches 292;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;

Qy 2058 AGATCAGAGTGGATGATCATATCATCATAGACAGATCAAGGCTTCGCTAGAGCTAC 2117
Db 160 AGATCAGAGTGGATGATCATATCATCATAGACAGATCAAGGCTTCGCTAGAGCTAC 219

QY	2118	AACAGAGGATTCAGGCAATTACCTCTGCCATGCGGTGAAACATGGGTTCTATACAAATC	2177
Db	220	AACAGAGGATTCAGGCAATTACCTCTGCCATGCGGTGAAACATGGGTTCTATACAAATC	279

Oy	2178	TTCTTAAGGTAACCTGGAGTCA TTGACA CAGAGCATTTGGAGAACTTCTCATTAAG	2237
Db	280	TTCTTAAGGTAACCTGGAGTCA TTGACA CAGAGCATTTGGAGAACTTCTCATTAAG	339

Qy 2238 ATGATGATGAGATGGCTCTTAAGACCAAGAATGTCCAAATGATGACCACTAGCCAGA 2297
340 ATGATGATGAGATGGCTCTTAAGACCAAGAATGTCCAAATGATGACCACTAGCCAGA 399

QY 2298 AGGTCTGTGACGAGACTTCATGCGCTCA1CACCACCCCAATCTCAACAC 2349
400 AGGTCTGTGACGAGACTTCATGCGCTCA1CACCACCCCAATCTCAACAC 451
db

RESULT 9
US-09-864-761-37943/-

```

: Sequence 27943, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
:

```

APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

```
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 27943
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g11421514, EVALUE 8.00e-89
OTHER INFORMATION: EST HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-09-864-761-27943

Query Match 5.9%; Score 160; DB 9; Length 172;
Best Local Similarity 100.0%; Pred. No. 6,7e-77;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 310 AGAATGTTGAATCAACAAATGATCATCTTTCATGAGCTTGGCCAAACAGCTCCAGTTA 369
Db 160 AGAATGTTGAATCAACAAATGATCATCTTTCATGAGCTTGGCCAAACAGCTCCAGTTA 101
Cy 370 TCATACCTCTCTTTGGATGAGGAGGAGTATGCTGATGTTGGAGCAAGATCAAT 429
Db 100 TCATACCTCTCTTTGGATGAGGAGGAGTATGCTGATGTTGGAGCAAGATCAAT 41
```

```

Cy 430 ATTTTCATTCGACCTGGTATATATCAAGATTTTCAAAAG 469
Db 40 ATTTTCATTCGACCTGGTATATATCAAGATTTTCAAAAG 1

RESULT 10
US-09-864-761-11363/c
Sequence 11363, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 11363
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-11363

Query Match 5.9%; Score 160; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 7.4e-77;
```


SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 28381
LENGTH: 123
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
OTHER INFORMATION: EST HUMAN HIT: BF667677.1, EVALUE 3.00e-63
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
OTHER INFORMATION: NT HIT: g11421514, EVALUE 2.00e-63
US-09-864-761-28381

Query Match 4.5%; Score 123; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. NO. 1.5e-56;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1937 CCTGAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCAGTCCG 1936
Db 1 CCTGAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCAGTCCG 60

Qy 1997 AAGTCGACAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGCGAAAGAA 2056
Db 61 AAGTCGACAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGCGAAAGAA 120

Qy 2057 GAG 2059
Db 121 GAG 123

RESULT 13
US-09-864-761-6056/c
Sequence 6056, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 6056
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
US-09-864-761-6056

Query Match 4.5%; Score 122; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. NO. 5.9e-56;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 532 GAAGAATGCTGCTATTTTCATCAAGTACTTAAGCATATATAGACTCACTTGTACGC 591
Db 407 GAAGAATGCTGCTATTTTCATCAAGTACTTAAGCATATATAGACTCACTTGTACGC 348

Qy 592 CTGTGAACGGGGCTTTTCATCAATTGGCACTTCACTTAATGAATGACATCACTCTGA 651
Db 347 CTGTGAACGGGGCTTTTCATCAATTGGCACTTCACTTAATGAATGACATCACTCTGA 288

Qy 652 GG 653
Db 287 GG 286

RESULT 14
US-09-864-761-22893/c
Sequence 22893, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 22893
LENGTH: 101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ACC04648.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: NT HIT: g11i21514, EVALUE 3.00e-50
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-09-864-761-22893

Query Match 3.7%; Score 101; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.8e-44;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 AAGATGTCATTAATTCATCAAGTACTTAAGGCATATATACAGACTGTTACGCC 592
DB 101 AAGATGTCATTAATTCATCAAGTACTTAAGGCATATATACAGACTGTTACGCC 42

QY 593 TGTGACGGGGGCTTTTCATCCCAATTGCACCTACATTGA 633
DB 41 TGTGACGGGGGCTTTTCATCCCAATTGCACCTACATTGA 1

RESULT 15
US-09-774-490-6
Sequence 6, Application US/09774490
Patent No. US20010034332A1

GENERAL INFORMATION:
APPLICANT: Jin, Shengfang

TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
FILE REFERENCE: 07334-138001
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: US 60/779,191
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6
LENGTH: 2137
TYPE: DNA

ORGANISM: Homo sapiens
US-09-774-490-6

Query Match 2.6%; Score 71; DB 9; Length 2137;

Best Local Similarity 100.0%; Pred. No. 7.8e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGTTAACAAGCTAGTAATGATGACCGACGCGGGGTGCG 60
DB 17 AATCTTTATTTATGATGTTAACAAGCTAGTAATGATGACCGACGCGGGGTGCG 76

QY 61 ACCCAGCGCTC 71
DB 77 ACCCAGCGCTC 87

RESULT 16
US-10-027-632-238542
Sequence 238542, Application US/10027632
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 238542

LENGTH: 520
TYPE: DNA
ORGANISM: Human

US-10-027-632-238542

Query Match 2.5%; Score 67; DB 13; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 AGACGACACAGACGACAGATATATAGAAATGAGAGACCCACTGACTCTGTTCAGACTTA 1909
DB 364 AGACGACACAGACGACAGATATATAGAAATGAGAGACCCACTGACTCTGTTCAGACTTA 423

QY 1910 CACCATG 1916
DB 424 CACCATG 430

RESULT 17
US-10-027-632-238543
Sequence 238543, Application US/10027632

GENERAL INFORMATION:
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29


```

; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 3299
; LENGTH: 115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)..(26)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (84)..(84)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (115)..(115)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3299

```

```

Query Match          0.8%; Score 22; DB 14; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 61 GTCGACCCACGCGTCGGGAGT 82

```

```

RESULT 25
US-09-910-943-236
; Sequence 236, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briuanlou, Ali
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 236
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(711)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-236

```

```

Query Match          0.8%; Score 22; DB 9; Length 711;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 48 GTCGACCCACGCGTCGGGAGT 69

```

```

RESULT 26
US-09-910-943-361
; Sequence 361, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briuanlou, Ali
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 361
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(748)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-361

```

```

Query Match          0.8%; Score 22; DB 9; Length 748;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 49 GTCGACCCACGCGTCGGGAGT 70

```

```

RESULT 27
US-09-897-006-6/c
; Sequence 6, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 5732
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-6

```

```

Query Match          0.8%; Score 22; DB 10; Length 5732;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AATCTTTATTTATCGATGTT 22
    |||||||
Db 5112 AATCTTTATTTATCGATGTT 5091

```

```

RESULT 28
US-09-897-511A-6/c
; Sequence 6, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors

```

FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 5732
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-6

Query Match 0.8%; Score 22; DB 11; Length 5732;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTT 22
|||||
Db 5112 AATCTTTATTTATGATGTT 5091

RESULT 29
US-08-786-531B-4/c
Sequence 4, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 7352 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-4

Query Match 0.8%; Score 22; DB 8; Length 7352;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTT 22
|||||
Db 4371 AATCTTTATTTATGATGTT 4350

RESULT 30
US-08-786-531B-1/c
Sequence 1, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-1

Query Match 0.8%; Score 22; DB 8; Length 7353;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTT 22
|||||
Db 4372 AATCTTTATTTATGATGTT 4351

RESULT 31
US-09-897-006-7/c
Sequence 7, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory


```

; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 9183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-7

Query Match      0.8%; Score 22; DB 10; Length 9183;
Best Local Similarity 100.0%; Pred.No. 0.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGTT 22
      |||||
Db      8563 AATCTTTATTTATCGATGTT 8542

RESULT 32
US-09-897-511A-7/c
; Sequence 7, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 9183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-7

Query Match      0.8%; Score 22; DB 11; Length 9183;
Best Local Similarity 100.0%; Pred.No. 0.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGTT 22
      |||||
Db      8563 AATCTTTATTTATCGATGTT 8542

RESULT 33
US-09-924-035A-528/c
; Sequence 528, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jim
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
```

```

; SEQ ID NO 528
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-528

Query Match      0.8%; Score 21; DB 10; Length 332;
Best Local Similarity 100.0%; Pred.No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
      |||||
Db      322 GTGACCCACGCGTCCGGAG 302

RESULT 34
US-09-925-300-828
; Sequence 828, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 828
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (21)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (128)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-828

Query Match      0.8%; Score 21; DB 10; Length 442;
Best Local Similarity 100.0%; Pred.No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
      |||||
Db      56 GTGACCCACGCGTCCGGAG 76
```

RESULT 35
US-09-960-352-4373
; Sequence 4373, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4373
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 19-LIB3058-036-Q1-K1-E3
US-09-960-352-4373

Query Match 0.8%; Score 21; DB 10; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CGACCCACCGCTCCGGAGTA 79
|||||
Db 6 CGACCCACCGCTCCGGAGTA 26

RESULT 36
US-10-027-632-235696/C
; Sequence 235696, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235696
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(690)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-235696

Query Match 0.8%; Score 21; DB 13; Length 690;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 AAGACATCTGAAAGATGTG 542
|||||

Db 513 AAGACATCTGAAAGATGTG 493

RESULT 37
US-10-106-698-1858
; Sequence 1858, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1858
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (13)-(13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (18)-(18)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (25)-(25)
; NAME/KEY: misc.feature
; LOCATION: (572)-(572)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1858

Query Match 0.8%; Score 21; DB 14; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGCACCCACCGCTCCGGAG 77
|||||
Db 76 GTGCACCCACCGCTCCGGAG 96

RESULT 38
US-09-910-943-67
; Sequence 67, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briylanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7539/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(725)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-67

Query Match 0.8%; Score 21; DB 9; Length 725;

Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 43 GTGACCCACGCGTCCGGAG 63

RESULT 39
US-09-910-943-448

; Sequence 448, Application US/09910943
; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briyvanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 448

; LENGTH: 730

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)_(730)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-448

Query Match 0.8%; Score 21; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 44 GTGACCCACGCGTCCGGAG 64

RESULT 40
US-09-910-943-674

; Sequence 674, Application US/09910943
; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briyvanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 674

; LENGTH: 745

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)_(745)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-674

Query Match 0.8%; Score 21; DB 9; Length 745;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 54 GTGACCCACGCGTCCGGAG 74

RESULT 41

US-09-910-943-359
; Sequence 359, Application US/09910943
; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briyvanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 359

; LENGTH: 749

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)_(749)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-359

Query Match 0.8%; Score 21; DB 9; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 55 GTGACCCACGCGTCCGGAG 75

RESULT 42
US-09-910-943-428

; Sequence 428, Application US/09910943
; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briyvanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

; CURRENT FILING DATE: 2001-07-23

; NUMBER OF SEQ ID NOS: 742

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 428

; LENGTH: 749

; TYPE: DNA

; ORGANISM: Xenopus laevis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)_(749)

; OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-428

Query Match 0.8%; Score 21; DB 9; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 50 GTGACCCACGCGTCCGGAG 70

RESULT 43
US-09-910-943-365

; Sequence 365, Application US/09910943
; Patent No. US20020081610A1

; GENERAL INFORMATION:

; APPLICANT: Hemmati-Briyvanlou, Ali

; APPLICANT: Altman, Curtis

; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

; FILE REFERENCE: 7529/IG148US1

; CURRENT APPLICATION NUMBER: US/09/910,943

/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 365
/ LENGTH: 750
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(750)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-365

Query Match 0.8%; Score 21; DB 9; Length 750;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 55 GTGACCCACGCGTCCGGAG 75

RESULT 44
US-09-910-943-504
/ Sequence 504, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briylanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 504
/ LENGTH: 750
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(750)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-504

Query Match 0.8%; Score 21; DB 9; Length 750;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 49 GTGACCCACGCGTCCGGAG 69

RESULT 45
US-09-910-943-383
/ Sequence 383, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briylanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 383
/ LENGTH: 781
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc feature

/ LOCATION: (1)..(781)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-383

Query Match 0.8%; Score 21; DB 9; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 43 GTGACCCACGCGTCCGGAG 63

RESULT 46
US-09-910-943-609
/ Sequence 609, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briylanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 609
/ LENGTH: 784
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(784)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-609

Query Match 0.8%; Score 21; DB 9; Length 784;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 68 GTGACCCACGCGTCCGGAG 88

RESULT 47
US-10-198-846-14063
/ Sequence 14063, Application US/10198846
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinhmann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14063
/ LENGTH: 883
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-198-846-14063

Query Match 0.8%; Score 21; DB 14; Length 883;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
DB 4 GTGACCCACGCGTCCGGAG 24

RESULT 48

US-10-037-270-33
; Sequence 33, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 33
LENGTH: 1181
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)..(874)
US-10-037-270-33

Query Match

Best Local Similarity 0.8%; Score 21; DB 14; Length 1181;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
DB 59 GTGACCCACGCGTCCGGAG 79

RESULT 49

US-10-106-698-1342
; Sequence 1342, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 1342
LENGTH: 1280
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (83)..(83)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (121)..(121)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (154)..(154)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (1166)..(1166)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1342

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
DB 243 GTGACCCACGCGTCCGGAG 263

RESULT 50

US-09-925-776-1
; Sequence 1, Application US/09925776
; Patent No. US20020038007A1
; GENERAL INFORMATION:
APPLICANT: AMES, ROBERT S., JR.
APPLICANT: SARAU, HENRY M.
APPLICANT: FOLEY, JAMES J.
APPLICANT: BERGSMAN, DEREK J.
APPLICANT: ELLIS, CATHERINE E.
TITLE OF INVENTION: A METHOD OF FINDING AGONIST AND
TITLE OF INVENTION: ANTAGONIST TO HUMAN 11CB SPLICED VARIANT
FILE REFERENCE: GP-50003-D2
CURRENT APPLICATION NUMBER: US/09/925,776
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/032,763
PRIOR FILING DATE: 1996-12-11
PRIOR APPLICATION NUMBER: 08/984,288
PRIOR FILING DATE: 1997-12-03
PRIOR APPLICATION NUMBER: 60/073,747
PRIOR FILING DATE: 1998-02-05
PRIOR APPLICATION NUMBER: 09/060,504
PRIOR FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1385
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-925-776-1

Query Match 0.8%; Score 21; DB 9; Length 1385;
Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
 |||||
 Db 50 GTGACCCACGCGTCCGGAG 70

RESULT 51
 US-10-119-926-44
 ; Sequence 44, Application US/10119926
 ; Publication No. US20030104413A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Xu, Chongjun
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Dimañac, Radoje T.
 ; TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and
 ; FILE REFERENCE: Polypeptides
 ; FILE REFERENCE: 789CIP2BCON
 ; CURRENT APPLICATION NUMBER: US/10/119,926
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/574,454
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/519,705
 ; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: pc_fl_genes Version 1.0
 ; SEQ ID NO 44
 ; LENGTH: 1394
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (415)..(1227)
 US-10-119-926-44

Query Match 0.8%; Score 21; DB 14; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
 |||||
 Db 33 GTGACCCACGCGTCCGGAG 53

RESULT 52
 US-10-037-270-664
 ; Sequence 664, Application US/10037270
 ; Publication No. US20030104529A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunting
 ; APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei
 APPLICANT: Tillinghast, John
 APPLICANT: Dimañac, Radoje T.
 TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 FILE REFERENCE: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/10/037,270
 CURRENT FILING DATE: 2002-01-04
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1104
 SOFTWARE: pc_fl_genes Version 1.0
 SEQ ID NO 664
 LENGTH: 1572
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (330)..(704)
 US-10-037-270-664

Query Match 0.8%; Score 21; DB 14; Length 1572;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
 |||||
 Db 280 GTGACCCACGCGTCCGGAG 300

RESULT 53
 US-09-838-573-1
 ; Sequence 1, Application US/09838573
 ; Patent No. US20020034783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel
 ; TITLE OF INVENTION: 39228, A NOVEL HUMAN ALCOHOL
 ; TITLE OF INVENTION: DEHYDROGENASE AND USES THEREFOR
 ; FILE REFERENCE: MNI-143
 ; CURRENT APPLICATION NUMBER: US/09/838,573
 ; CURRENT FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/197,747
 ; PRIOR FILING DATE: 2000-04-18
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1808
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (285)..(1418)
 US-09-838-573-1

Query Match 0.8%; Score 21; DB 9; Length 1808;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
 |||||
 Db 4 GTGACCCACGCGTCCGGAG 24

RESULT 54
 US-10-160-501-1
 ; Sequence 1, Application US/10160501
 ; Publication No. US2003005919A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel E.
 ; APPLICANT: Rudolph-Owen, Laura A.
 ; APPLICANT: Kapeller-Libermann, Rosana

```

; TITLE OF INVENTION: NOVEL HUMAN 39228, 21956, 25856, 22244, 8701, 32263,
; TITLE OF INVENTION: 50250, 55158, 47765, 62088, 50566, AND 48118
; TITLE OF INVENTION: MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-250
; CURRENT APPLICATION NUMBER: US/10/160,501
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 09/838,573
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/197,747
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/870,133
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,649
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/870,130
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,640
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/862,535
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,961
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,383
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,506
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/860,821
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,449
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,110
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,650
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/907,509
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,385
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/945,327
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,425
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/318,581
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FASTSEQ Version 4.0
; SEQ ID NO 1
; LENGTH: 1808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (285)...(1418)
; US-10-160-501-1

Query Match      0.8%; Score 21; DB 14; Length 1808;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      57  GTGACCCACGCGTCGGGAG 77
      |||||
Db      4  GTGACCCACGCGTCGGGAG 24

```

```

RESULT 55
US-09-925-302-192
; Sequence 192, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302

```

```

; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 192
; LENGTH: 2570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2561)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-302-192

Query Match      0.8%; Score 21; DB 9; Length 2570;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      58  TCGACCCACGCGTCGGGAGT 78
      |||||
Db      1  TCGACCCACGCGTCGGGAGT 21

```

```

RESULT 56
US-09-796-753-89
; Sequence 89, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002

```



```

; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 89
; LENGTH: 2810
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)..(1527)
; US-09-796-753-89

Query Match          0.8%; Score 21; DB 11; Length 2810;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
Db      1 GTGACCCACGCGTCCGGAG 21

RESULT 57
US-09-782-980-67
; Sequence 67, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIPE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STWST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
```

```

; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/672,721
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/049,799
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 2815
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (197)..(2215)
; US-09-782-980-67

Query Match          0.8%; Score 21; DB 9; Length 2815;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
Db      1 GTGACCCACGCGTCCGGAG 21

RESULT 58
US-09-897-006-4/c
; Sequence 4, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 4207
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-09-897-006-4

Query Match          0.8%; Score 21; DB 10; Length 4207;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCATGT 21
Db      3587 AATCTTTATTTATCATGT 3567

RESULT 59
US-09-897-511A-4/c
; Sequence 4, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
```

NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 4207
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-4

Query Match 0.8%; Score 21; DB 11; Length 4207;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3587 AATCTTTATTTATCGATGT 3567

RESULT 60
US-09-897-006-5/c
Sequence 5, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Expression Vectors
FILE REFERENCE: GALA-06415
CURRENT APPLICATION NUMBER: US/09/897,006
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,851
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 4210
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-006-5

Query Match 0.8%; Score 21; DB 10; Length 4210;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3595 AATCTTTATTTATCGATGT 3575

RESULT 61
US-09-897-511A-5/c
Sequence 5, Application US/09897511A
Publication No. US20030092882A1
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Miller, Linda
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 4210
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-5

Query Match 0.8%; Score 21; DB 11; Length 4210;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3595 AATCTTTATTTATCGATGT 3575

RESULT 62
US-09-897-006-13/c
Sequence 13, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Expression Vectors
FILE REFERENCE: GALA-06415
CURRENT APPLICATION NUMBER: US/09/897,006
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,851
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 6255
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-006-13

Query Match 0.8%; Score 21; DB 10; Length 6255;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 5635 AATCTTTATTTATCGATGT 5615

RESULT 63
US-09-897-511A-13/c
Sequence 13, Application US/09897511A
Publication No. US20030092882A1
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Miller, Linda
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 6255
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-13

Query Match 0.8%; Score 21; DB 11; Length 6255;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 5635 AATCTTTATTTATCGATGT 5615

```
RESULT 64
US-09-897-006-34/c
; Sequence 34, Application US/09897006
; Patent No. US20020106723A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 9511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-34

Query Match          0.8%; Score 21; DB 10; Length 9511;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGT 21
        |||||
Db      6526 AATCTTTATTTATCGATGT 6506

RESULT 65
US-09-897-511A-34/c
; Sequence 34, Application US/09897511A
; Publication No. US20030092862A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 9511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-34

Query Match          0.8%; Score 21; DB 11; Length 9511;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGT 21
        |||||
Db      6526 AATCTTTATTTATCGATGT 6506

RESULT 66
US-10-106-698-4090
; Sequence 4090, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005PI
```

```
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4090
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (49)..(49)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (62)..(62)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (88)..(88)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (93)..(93)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-4090

Query Match          0.7%; Score 20; DB 14; Length 100;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTCGACCCAGCGGTCGGGA 76
        |||||
Db      24 GTCGACCCAGCGGTCGGGA 43

RESULT 67
US-09-764-891-522
; Sequence 522, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 522
; LENGTH: 122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-522

Query Match          0.7%; Score 20; DB 11; Length 122;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTCGACCCAGCGGTCGGGA 76
        |||||
Db      31 GTCGACCCAGCGGTCGGGA 50
```

```

PRIORITY FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 2365
LENGTH: 197
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15)..(15)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (21)..(22)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (27)..(27)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (41)..(41)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (192)..(192)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (194)..(197)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2365

Query Match          0.7%; Score 20; DB 14; Length 197;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57  GTGACCCACGCGTCGCGGA 76
      |||||
DB      125  GTGACCCACGCGTCGCGGA 144

RESULT 70
US-09-918-995-8601
Sequence 8601, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIORITY FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8601
LENGTH: 327
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(327)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8601

Query Match          0.7%; Score 20; DB 11; Length 327;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57  GTGACCCACGCGTCGCGGA 76
      |||||
DB      32  GTGACCCACGCGTCGCGGA 51

RESULT 71
US-09-986-480-113

```

```
/ Sequence 113, Application US/09986480
/ Publication No. US20030027999A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 143 Human Secreted Proteins
/ FILE REFERENCE: PS500P1
/ CURRENT APPLICATION NUMBER: US/09/986,480
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: PCT/US00/12788
/ PRIOR FILING DATE: 2000-05-11
/ PRIOR APPLICATION NUMBER: US 60/134,068
/ PRIOR FILING DATE: 1999-05-13
/ NUMBER OF SEQ ID NOS: 456
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 113
/ LENGTH: 355
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-986-480-113

Query Match
Best Local Similarity 100.0%; Score 20; DB 11; Length 355;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGGCTCCGGGA 76
DB 5 GTCGACCCACGGCTCCGGGA 24

RESULT 72
US-10-198-846-10302/c
/ Sequence 10302, Application US/10198846
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinhann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10302
/ LENGTH: 367
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-198-846-10302

Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 367;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGGCTCCGGGA 76
DB 362 GTCGACCCACGGCTCCGGGA 343

RESULT 73
US-09-925-300-802/c
/ Sequence 802, Application US/09925300
/ Patent No. US20020151681A1
/ GENERAL INFORMATION:
/ APPLICANT: Craig Rosen,
/ APPLICANT: Steve Ruben,
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antipodides
/ FILE REFERENCE: PA101
/ CURRENT APPLICATION NUMBER: US/09/925,300
```

```
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05988
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1890
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 802
/ LENGTH: 402
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (149)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (310)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (322)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (344)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (363)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (383)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-802

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 402;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGGCTCCGGGA 76
DB 41 GTCGACCCACGGCTCCGGGA 22

RESULT 74
US-10-106-698-2172
/ Sequence 2172, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 2172
/ LENGTH: 405
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (24)-(24)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (357)-(357)
/ OTHER INFORMATION: n equals a,t,g, or c
```

NAME/KEY: misc.feature
LOCATION: (385)..(385)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (394)..(396)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2172

Query Match 0.7%; Score 20; DB 14; Length 405;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCGGGA 76
Db 72 GTCGACCCACGCGTCGGGA 91

RESULT 75
US-09-960-352-13318
Sequence 13318, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 13318
LENGTH: 430
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 57-LIB3057-004-Q1-K1-G2
US-09-960-352-13318

Query Match 0.7%; Score 20; DB 10; Length 430;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 ACCCAGCGGTCGCGGAGTAG 80
Db 1 ACCCAGCGGTCGCGGAGTAG 20

Search completed: August 1, 2003, 02:13:41
Job time : 602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 31, 2003, 19:37:11 ; Search time 7225 Seconds

(without alignments)
12286.241 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aactcttattatcgatg.....aggtcttttctcctaataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32997241 seqs, 1638322548 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/1/pna/PCRTUS_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCRTUS_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US092A_COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	US-09-774-490-1	Sequence 1, Appl
2	2517	92.9	2848	US-10-247-671-83	Sequence 83, Appl
3	2517	92.9	2848	US-60-172-373-8935	Sequence 8935, Ap
4	2517	92.9	2848	US-60-323-784-83	Sequence 83, Appl

5	2508	92.6	2530	1	PCT-US02-07826-283	Sequence 283, App
6	2508	92.6	2530	15	US-09-053-375B-1350	Sequence 1350, App
7	2508	92.6	2530	45	US-10-097-340-283	Sequence 283, App
8	2508	92.6	2530	49	US-10-262-538-9	Sequence 9, App1
9	2508	92.6	2601	9	US-08-488-935-1	Sequence 1, App1
10	2508	92.6	2601	9	US-08-489-057-1	Sequence 1, App1
11	2508	92.6	2601	45	US-10-067-633-53	Sequence 53, App1
12	2226	82.2	2857	47	US-10-170-235-19890	Sequence 19890, App
13	1388	51.2	2256	28	US-09-652-164-8352	Sequence 8352, App
14	1388	51.2	2256	28	US-09-652-816-7308	Sequence 7308, App
15	1388	51.2	2256	28	US-09-652-918-7516	Sequence 7516, App
16	817	30.2	944	32	US-09-724-676-18078	Sequence 18078, App
17	817	30.2	944	32	US-09-724-676-18078	Sequence 18078, App
18	405	15.0	456	36	US-09-864-761-30992	Sequence 30992, App
19	405	15.0	456	48	US-10-203-135-19786	Sequence 19786, App
20	405	15.0	456	48	US-10-203-135-20203	Sequence 20203, App
21	379	14.0	379	17	US-09-205-070-45112	Sequence 45112, App
22	379	14.0	379	19	US-09-321-214-10943	Sequence 10943, App
23	379	14.0	379	19	US-09-340-623-45112	Sequence 45112, App
24	379	14.0	379	22	US-09-516-335-10943	Sequence 10943, App
25	379	14.0	379	32	US-09-733-811-10943	Sequence 10943, App
26	379	14.0	379	32	US-09-733-811A-10943	Sequence 10943, App
27	379	14.0	379	37	US-09-898-888-45112	Sequence 45112, App
28	379	14.0	379	37	US-09-898-888A-45112	Sequence 45112, App
29	379	14.0	379	42	US-09-975-640-10943	Sequence 10943, App
30	379	14.0	379	42	US-09-975-640A-10943	Sequence 10943, App
31	370	13.7	497	23	US-09-534-856-4	Sequence 4, App1
32	303	11.2	354	1	PCT-US02-11475A-736	Sequence 736, App
33	303	11.2	354	26	US-09-609-448A-736	Sequence 736, App
34	303	11.2	354	28	US-09-649-811-736	Sequence 736, App
35	303	11.2	354	35	US-09-833-263-736	Sequence 736, App
36	303	11.2	354	39	US-09-922-217-736	Sequence 736, App
37	303	11.2	354	44	US-10-025-380-736	Sequence 736, App
38	292	10.8	496	36	US-09-864-761-14444	Sequence 14444, App
39	292	10.8	496	48	US-10-203-135-7281	Sequence 7281, App
40	292	10.8	496	48	US-10-203-136-7203	Sequence 7203, App
41	256	9.4	427	20	US-09-399-720-14058	Sequence 14058, App
42	256	9.4	427	39	US-09-921-378-14058	Sequence 14058, App
43	254	9.4	254	21	US-09-442-385-770	Sequence 770, App
44	228	8.4	268	34	US-09-824-130-5891	Sequence 5891, App
45	169	6.2	636	70	US-09-160-203-79	Sequence 79, App1
46	161	5.9	636	70	US-09-169-840-156	Sequence 156, App
47	161	5.9	164	44	US-10-029-386-18184	Sequence 18184, App
48	161	5.9	591	44	US-10-029-386-4484	Sequence 4484, App
49	160	5.9	172	1	PCT-US01-00663-20125	Sequence 20125, App
50	160	5.9	172	2	PCT-US01-00663-20125	Sequence 20125, App
51	160	5.9	172	36	US-09-864-761-27943	Sequence 27943, App
52	160	5.9	172	47	US-10-182-993-19495	Sequence 19495, App
53	160	5.9	172	48	US-10-203-134-20052	Sequence 20052, App
54	160	5.9	172	48	US-10-203-135-19754	Sequence 19754, App
55	160	5.9	172	48	US-10-203-136-20171	Sequence 20171, App
56	160	5.9	172	48	US-10-203-137-20125	Sequence 20125, App
57	160	5.9	172	48	US-10-203-139-19510	Sequence 19510, App
58	160	5.9	191	44	US-10-029-386-16491	Sequence 16491, App
59	160	5.9	484	1	PCT-US01-00663-7072	Sequence 7072, App
60	160	5.9	484	2	PCT-US01-00663-7072	Sequence 7072, App
61	160	5.9	484	36	US-09-864-761-11363	Sequence 11363, App
62	160	5.9	484	47	US-10-182-993-6833	Sequence 6833, App
63	160	5.9	484	48	US-10-203-134-7118	Sequence 7118, App
64	160	5.9	484	48	US-10-203-135-7342	Sequence 7342, App
65	160	5.9	484	48	US-10-203-136-7262	Sequence 7262, App
66	160	5.9	484	48	US-10-203-137-7012	Sequence 7012, App
67	160	5.9	484	48	US-10-203-139-7012	Sequence 7012, App
68	160	5.9	526	44	US-10-029-386-2791	Sequence 2791, App
69	144	5.3	446	36	US-09-864-761-11810	Sequence 11810, App
70	144	5.3	446	47	US-10-182-993-7713	Sequence 7713, App
71	144	5.3	446	48	US-10-203-135-8417	Sequence 8417, App
72	144	5.3	446	48	US-10-203-136-8331	Sequence 8331, App
73	144	5.3	574	70	US-09-160-203-771	Sequence 771, App
74	144	5.3	728	70	US-09-160-203-838	Sequence 838, App
75	123	4.5	123	36	US-09-864-761-28381	Sequence 28381, App
76	123	4.5	123	47	US-10-182-993-20377	Sequence 20377, App
77	123	4.5	123	48	US-10-203-135-20831	Sequence 20831, App

78	123	4.5	123	48	US-10-203-136-21236	Sequence 21236, App
79	122	4.5	456	1	PCT-US01-00663-6904	Sequence 6904, App
80	122	4.5	456	2	PCT-US01-00663-6904	Sequence 6904, App
81	122	4.5	456	36	US-09-864-761-6056	Sequence 6056, App
82	122	4.5	456	47	US-10-182-993-6667	Sequence 6667, App
83	122	4.5	456	47	US-10-182-993-6056	Sequence 6056, App
84	122	4.5	456	48	US-10-203-134-7929	Sequence 7929, App
85	122	4.5	456	48	US-10-203-135-7146	Sequence 7146, App
86	122	4.5	456	48	US-10-203-136-7070	Sequence 7070, App
87	122	4.5	456	48	US-10-203-137-6904	Sequence 6904, App
88	122	4.5	456	48	US-10-203-139-6862	Sequence 6862, App
89	122	4.5	556	72	US-09-181-996-169	Sequence 169, App
90	115	4.2	588	71	US-09-177-571-733	Sequence 733, App
91	109	4.0	405	80	US-06-252-833-39848	Sequence 39848, App
92	103	3.8	385	71	US-06-177-571-962	Sequence 962, App
93	103	3.8	385	71	US-06-177-571-963	Sequence 963, App
94	101	3.7	101	1	PCT-US01-00663-20219	Sequence 20219, App
95	101	3.7	101	2	PCT-US01-00663-20219	Sequence 20219, App
96	101	3.7	101	36	US-09-864-761-22893	Sequence 22893, App
97	101	3.7	101	47	US-10-182-993-19584	Sequence 19584, App
98	101	3.7	101	47	US-10-182-995-16039	Sequence 16039, App
99	101	3.7	101	48	US-10-203-134-20153	Sequence 20153, App
100	101	3.7	101	48	US-10-203-135-19859	Sequence 19859, App

ALIGNMENTS

RESULT 1	
US-09-774-490-1	
Sequence 1, Application US/09774490	
GENERAL INFORMATION:	
APPLICANT: Jin, Shengfang	
TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF	
FILE REFERENCE: 07334-138001	
CURRENT APPLICATION NUMBER: US/09/774,490	
PRIOR FILING DATE: 2001-01-31	
PRIOR APPLICATION NUMBER: US 60/179,191	
PRIOR FILING DATE: 2000-01-31	
NUMBER OF SEQ ID NOS: 6	
SOFTWARE: FastSeq for Windows, Version 4.0	
SEQ ID NO 1	
LENGTH: 2709	
TYPE: DNA	
ORGANISM: Mus musculus	
US-09-774-490-1	
Query Match	100.0%; Score 2709; DB 33; Length 2709;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2709; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	AATCTTATTTATGATGTTAACAAGCTAGTAATGATGCCAGCGGAGGTCG 60
1	AATCTTATTTATGATGTTAACAAGCTAGTAATGATGCCAGCGGAGGTCG 60
61	ACCAAGCGCTCGGAGTAGTGGCTGCTGCTCCATTGTCAGCCAGTCTATT 120
61	ACCAAGCGCTCGGAGTAGTGGCTGCTGCTCCATTGTCAGCCAGTCTATT 120
121	CCAGTTGTTGAACTTCTTCGCGGCACATATACAGAGAAGACTTAAGCAAG 180
121	CCAGTTGTTGAACTTCTTCGCGGCACATATACAGAGAAGACTTAAGCAAG 180
181	GACCTAGCGCTCGGAGTAGTGGCTGTTAAGTATGCTCTTTCTGGGGAGT 240
181	GACCTAGCGCTCGGAGTAGTGGCTGTTAAGTATGCTCTTTCTGGGGAGT 240
241	ATTACTTACAGCAAGCAACTATCAGATGGAGAGACATGTGCCAGGCTGAATT 300
241	ATTACTTACAGCAAGCAACTATCAGATGGAGAGACATGTGCCAGGCTGAATT 300
301	ATCTCAAGAAATGTTGATCCAAATGATGATCTTCAATGGCTTGGCAAG 360
301	ATCTCAAGAAATGTTGATCCAAATGATGATCTTCAATGGCTTGGCAAG 360

Db 301 ATCTTAAGAAAGTGTGGAATCCAAACATGTGATCACTTTCATGTGCGCAACAG 360
Qy 361 CTCGATTATCAATACCTTCCTTTGGATGAGAACCGAGTAGCTGTATGTTGAGCAAA 420
Db 361 CTCGATTATCAATACCTTCCTTTGGATGAGAACCGAGTAGCTGTATGTTGAGCAAA 420
Qy 421 GGATCAATATTTTCATTCGACCTGTGTTAATCAAGATTTTCAAAAGATTGTGTGCC 480
Db 421 GGATCAATATTTTCATTCGACCTGTGTTAATCAAGATTTTCAAAAGATTGTGTGCC 480
Qy 481 AGTATCTTAACAGAAAGATGATGAGAGTGGGTGAGAAAGACATCTCTAAAGATG 540
Db 481 AGTATCTTAACAGAAAGATGATGAGAGTGGGTGAGAAAGACATCTCTAAAGATG 540
Qy 541 TGTCTAATTTTCATCAAGTACTTAAGCATATATCACTCACTTGTAGCGCTGTGAAAC 600
Db 541 TGTCTAATTTTCATCAAGTACTTAAGCATATATCACTCACTTGTAGCGCTGTGAAAC 600
Qy 601 GGGGCTTTTCATCCAAATTTGCACTCACTGAAATTTGACATCACTTGAAGACATAT 660
Db 601 GGGGCTTTTCATCCAAATTTGCACTCACTGAAATTTGACATCACTTGAAGACATAT 660
Qy 661 TTTTAAAGCTGAGAACTCACTTTTGAAGAAAGCGCGTGGAGAGTCCATATGACCTTA 720
Db 661 TTTTAAAGCTGAGAACTCACTTTTGAAGAAAGCGCGTGGAGAGTCCATATGACCTTA 720
Qy 721 GCTGCTGACAGCATCCCTTTTAAATGATGAGAAATTAATCTCTGGAAGTGCAGTGT 780
Db 721 GCTGCTGACAGCATCCCTTTTAAATGATGAGAAATTAATCTCTGGAAGTGCAGTGT 780
Qy 781 TATGGGGGAGAACTTTGCTATCTTCGAACTCTTGGGACCAACCACTGAGACAGA 840
Db 781 TATGGGGGAGAACTTTGCTATCTTCGAACTCTTGGGACCAACCACTGAGACAGA 840
Qy 841 GAGAGATGATTTCCAGTGTCTCAATGATGCAAGTTCATTAAGTCCACCTCATCTCAGA 900
Db 841 GAGAGATGATTTCCAGTGTCTCAATGATGCAAGTTCATTAAGTCCACCTCATCTCAGA 900
Qy 901 GAGTACAACTCTGAAGATGACAAAGTATCTTTTCTCGTGAAGAAATGCAATGATG 960
Db 901 GAGTACAACTCTGAAGATGACAAAGTATCTTTTCTCGTGAAGAAATGCAATGATG 960
Qy 961 AGAACACTCTGGAAGAACTACTCAGCTGATGAAATAGTCAATGCAAGATGACTTTGG 1020
Db 961 AGAACACTCTGGAAGAACTACTCAGCTGATGAAATAGTCAATGCAAGATGACTTTGG 1020
Qy 1021 AGGGAACAGAACTGTGGAATTAATGCAACATTTCTGAAAGCTCGTGTGATTTGCTC 1080
Db 1021 AGGGAACAGAACTGTGGAATTAATGCAACATTTCTGAAAGCTCGTGTGATTTGCTC 1080
Qy 1081 AGTGCAGAGTCCAAATGATGACATCTTTGATGAACTGCAAGATGATTTGCTAAT 1140
Db 1081 AGTGCAGAGTCCAAATGATGACATCTTTGATGAACTGCAAGATGATTTGCTAAT 1140
Qy 1141 GAACTTTAAAGATCTTAAAGATCCAGTGTATGAGAGTGTTCAGACTTCCAGTAAGAT 1200
Db 1141 GAACTTTAAAGATCTTAAAGATCCAGTGTATGAGAGTGTTCAGACTTCCAGTAAGAT 1200
Qy 1201 TTTTCAAGGATGAGCGGTGTGTATGATGATGAGTGTGAGAGAGGTGTTCTTTGG 1260
Db 1201 TTTTCAAGGATGAGCGGTGTGTATGATGATGAGTGTGAGAGAGGTGTTCTTTGG 1260
Qy 1261 TCCATATGCCCAAGAGATGAGCCCACTATGATGAGAGTGTTCATCAAGAGAGTCCC 1320
Db 1261 TCCATATGCCCAAGAGATGAGCCCACTATGATGAGAGTGTTCATCAAGAGAGTCCC 1320
Qy 1321 CTATCCAGGCGCAGAACTTGTCCAGCAAAACATTTGTGTGTTTGACTCTAACAAGAA 1380
Db 1321 CTATCCAGGCGCAGAACTTGTCCAGCAAAACATTTGTGTGTTTGACTCTAACAAGAA 1380
Qy 1381 CCTTCTGTATGATGTTATTAACCTTTGCAAGAAAGTCACTCAGCAGTACATCAAGTGT 1440
Db 1381 CCTTCTGTATGATGTTATTAACCTTTGCAAGAAAGTCACTCAGCAGTACATCAAGTGT 1440

Qy 1441 TCTATGAACAATGCGCCCAATAGTATGATCAAAACGATGTAAATTAATTAATTAACAAT 1500
Db 1441 TCTATGAACAATGCGCCCAATAGTATGATCAAAACGATGTAAATTAATTAATTAACAAT 1500
Qy 1501 TGTGTGAGACCGAGTGTGAGAGAAAGATGACAGTATGATGATGTTATGGAACAGA 1560
Db 1501 TGTGTGAGACCGAGTGTGAGAGAAAGATGACAGTATGATGATGTTATGGAACAGA 1560
Qy 1561 TGTGTGAGACCGTCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 1561 TGTGTGAGACCGTCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1620
Qy 1621 GGTTCGTGGAAGAAATGACAGTGTTCGGGAAACCGACTCTATTTTCAAGCAATGAGCT 1680
Db 1621 GGTTCGTGGAAGAAATGACAGTGTTCGGGAAACCGACTCTATTTTCAAGCAATGAGCT 1680
Qy 1681 TTTCACTAAGCAGCAACAATATATATGTTCAACGCGTGGGTTGCGCAGCTCCCTTT 1740
Db 1681 TTTCACTAAGCAGCAACAATATATATGTTCAACGCGTGGGTTGCGCAGCTCCCTTT 1740
Qy 1741 AACCGGTGTATATTTACGGGAAAGCGTGTGAGAGTGTGCTGCGCCGAGACCTTTA 1800
Db 1741 AACCGGTGTATATTTACGGGAAAGCGTGTGAGAGTGTGCTGCGCCGAGACCTTTA 1800
Qy 1801 CTGTGCTTGGGATGATCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAG 1860
Db 1801 CTGTGCTTGGGATGATCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAG 1860
Qy 1861 ACGACAGATATTAAGAAATGAGAACCACTACCTACTGTTTCACTTACCAATGATTA 1920
Db 1861 ACGACAGATATTAAGAAATGAGAACCACTACCTACTGTTTCACTTACCAATGATTA 1920
Qy 1921 TCACTATGCGCACGCGCTGAAAGAGAAATCATATGATGATGATGATGATGATGATG 1980
Db 1921 TCACTATGCGCACGCGCTGAAAGAGAAATCATATGATGATGATGATGATGATGATG 1980
Qy 1981 TTTGGAATGACAGTCCGAAGTGCAGAGAGCGCTGATTTGGAATTCAGAGGCGAA 2040
Db 1981 TTTGGAATGACAGTCCGAAGTGCAGAGAGCGCTGATTTGGAATTCAGAGGCGAA 2040
Qy 2041 TGAAGAGGAAAGAGAGATCAGATGATGATCATATCATCAAGACAGATCAAGGCT 2100
Db 2041 TGAAGAGGAAAGAGAGATCAGATGATGATCATATCATCAAGACAGATCAAGGCT 2100
Qy 2101 TCTGTATGATGATCTAACAAGAGATTCAGGCAATTAATCTTGTGCAATGCGGTGAA 2160
Db 2101 TCTGTATGATGATCTAACAAGAGATTCAGGCAATTAATCTTGTGCAATGCGGTGAA 2160
Qy 2161 TGGGTTCATACAACTCTTTAAGTAACTGGAAGTCAATGACACAGACATTTGGA 2220
Db 2161 TGGGTTCATACAACTCTTTAAGTAACTGGAAGTCAATGACACAGACATTTGGA 2220
Qy 2221 AGAATCTTTCATTAAGATGATGAGATGAGTCTTAAGACCAAGAAATGTCATATG 2280
Db 2221 AGAATCTTTCATTAAGATGATGAGATGAGTCTTAAGACCAAGAAATGTCATATG 2280
Qy 2281 CATGACCTTACGCAAGAGTCTGTGACAGAGCTTCAATGACATCAACACCCCA 2340
Db 2281 CATGACCTTACGCAAGAGTCTGTGACAGAGCTTCAATGACATCAACACCCCA 2340
Qy 2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Db 2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Qy 2401 GCAAGGCGCAGAGATACCCAGAGAAACATTAAGAAATGAGCACTTAACAAGAAATTA 2460
Db 2401 GCAAGGCGCAGAGATACCCAGAGAAACATTAAGAAATGAGCACTTAACAAGAAATTA 2460
Qy 2461 GAAAGGTGAGAAACAGAGAGACCCAGATTTGAGAGGCGCACCCAGAGTGTCTGAGCTGC 2520
Db 2461 GAAAGGTGAGAAACAGAGAGACCCAGATTTGAGAGGCGCACCCAGAGTGTCTGAGCTGC 2520

QY 2521 ATTACCTTGAAACCTCAACAGTAGAACTTGCTAGACATAATCGGAAAAA 2580
DB 2521 ATTACCTTGAAACCTCAACAGTAGAACTTGCTAGACATAATCGGAAAAA 2580
QY 2581 TCGAATATACATGACCTTTTTCATGCGATTATGCGATTTTCAATGGTGGAAATTC 2640
DB 2581 TCGAATATACATGACCTTTTTCATGCGATTATGCGATTTTCAATGGTGGAAATTC 2640
QY 2641 AGCTGAGTTCCACCAATTAATTAATTCATGATTAATCTTCTTAATAGCTTTTTC 2700
DB 2641 AGCTGAGTTCCACCAATTAATTAATTCATGATTAATCTTCTTAATAGCTTTTTC 2700
QY 2701 CCTAATACC 2709
DB 2701 CCTAATACC 2709

RESULT 2
US-10-247-671-83
Sequence 83, Application US/10247671
GENERAL INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
PRIOR FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 186
SOFTWARE: PERL Program
SEQ ID NO 83
LENGTH: 2848
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1930967CBL
US-10-247-671-83

Query Match 92.9%; Score 2517; DB 48; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGCGACATGGGCTGGTTAATAGATTTGCTGCTCTTTCTGGGGAGTATTAATTACAC 252
DB 9 CTGCGACATGGGCTGGTTAATAGATTTGCTGCTCTTTCTGGGGAGTATTAATTACAC 68
QY 251 AAGAGCAACTATCAGATAGGAGAAACAATGTGCCAAGGCTGAATTAATCTCAAAAGA 312
DB 69 AAGAGCAACTATCAGATAGGAGAAACAATGTGCCAAGGCTGAATTAATCTCAAAAGA 128
QY 313 AATGTTGATCCACATATGTGATCACTTTCAATGCGCTTGCCACACAGCTCAATTATCA 372
DB 129 AATGTTGATCCACATATGTGATCACTTTCAATGCGCTTGCCACACAGCTCAATTATCA 188
QY 373 TACCTTCCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGAGATCACATATT 432
DB 189 TACCTTCCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAGAGATCACATATT 248
QY 433 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATTTGTTGCGCAGTATCTTACAC 492
DB 249 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATTTGTTGCGCAGTATCTTACAC 308
QY 493 CAGAAAGATGATGACAGTGGGCTGGAAGAAAGACATCTGGAAGAAATGTGCTAATTCAT 552
DB 309 CAGAAAGATGATGACAGTGGGCTGGAAGAAAGACATCTGGAAGAAATGTGCTAATTCAT 368
QY 553 CAGAGTACTTAAGGATATATATCAGACTATTTGTCGCTGTGGAACGGGGGCTTTTCA 612
DB 369 CAGAGTACTTAAGGATATATATCAGACTATTTGTCGCTGTGGAACGGGGGCTTTTCA 428

QY 613 TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAACTGGA 672
DB 429 TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAACTGGA 488
QY 673 GAATCTACATTTTGAAGAACGGCGCTGGAGAGATTCATATGACCTTAAGCTCTGACAGC 732
DB 489 GAATCTACATTTTGAAGAACGGCGCTGGAGAGATTCATATGACCTTAAGCTCTGACAGC 548
QY 733 ATCCCTTTTAATAGATGAGAAATTAATTAATCTGGAATCTGAGCTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTTAATAGATGAGAAATTAATTAATCTGGAATCTGAGCTGATTTTATGGGCGAGA 608
QY 793 CTTTGCTATCTTCCGAACTTTGGGCAACACCAATCAGACAGAGCATGATTC 852
DB 609 CTTTGCTATCTTCCGAACTTTGGGCAACACCAATCAGACAGAGCATGATTC 668
QY 853 CAGTGGCTCAATGATCCAAAGTTTATGAGTCCCACTCATCTGAGAGTGAACAATCC 912
DB 669 CAGTGGCTCAATGATCCAAAGTTTATGAGTCCCACTCATCTGAGAGTGAACAATCC 728
QY 913 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAAATGCAATAGATGAGAAACCTGCG 972
DB 729 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAAATGCAATAGATGAGAAACCTGCG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGGTCAATATGCAAGATGACCTTTGAGGGGCAAGAAG 1032
DB 789 AAAAGCTACTCAGCTAGAAATAGGTCAATATGCAAGATGACCTTTGAGGGGCAAGAAG 848
QY 1033 TCTGATGAATTAATGAGCAACATTCCTCAAGGCTGTGATTTGCTGATGAGTCCAGGCTCC 1092
DB 849 TCTGATGAATTAATGAGCAACATTCCTCAAGGCTGTGATTTGCTGATGAGTCCAGGCTCC 908
QY 1093 AAATGACATGACATCTATTTTGAATGAGTCAAGATGATTAATCTTAATGAATCTTAAGA 1152
DB 909 AAATGACATGACATCTATTTTGAATGAGTCAAGATGATTAATCTTAATGAATCTTAAGA 968
QY 1153 TCTTAATAATCCAGTTGTATATGAGTGTATTAACAATTCAGATTAATTTTCAAGGATC 1212
DB 969 TCTTAATAATCCAGTTGTATATGAGTGTATTAACAATTCAGATTAATTTTCAAGGATC 1028
QY 1213 AGCCGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272
DB 1029 AGCCGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1088
QY 1273 CAGGATGAGACCAACTATCAATGAGTCTTATCAAGAAAGATCCCTATCCAGGCTC 1332
DB 1089 CAGGATGAGACCAACTATCAATGAGTCTTATCAAGAAAGATCCCTATCCAGGCTC 1148
QY 1333 AGGAATTTGCTCCAGCAAAACATTTGATTTGATCTTAACAAGACCTTCTGATGA 1392
DB 1149 AGGAATTTGCTCCAGCAAAACATTTGATTTGATCTTAACAAGACCTTCTGATGA 1208
QY 1393 TGTATTAACCTTTTGAAGAGATCCAGCATGATGATGATGATGATGATGATGATGATGATG 1452
DB 1209 TGTATTAACCTTTTGAAGAGATCCAGCATGATGATGATGATGATGATGATGATGATGATG 1268
QY 1453 TCGCCCAATAGATCAAAAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
DB 1269 TCGCCCAATAGATCAAAAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1328
QY 1513 AGTGAATCAGAAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1572
DB 1329 AGTGAATCAGAAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1388
QY 1573 TCTTAAGTATGTTCAATCTTAAGAGACCTTGTATGATTAATTAAGAGGTTTCTGCTGGA 1632
DB 1389 TCTTAAGTATGTTCAATCTTAAGAGACCTTGTATGATTAATTAAGAGGTTTCTGCTGGA 1448
QY 1633 AGAATGACAGTTTGGGAAACGAGCTGATTTTCAAGATGAGAGCTTCACTAAGCA 1692
DB 1449 AGAATGACAGTTTGGGAAACGAGCTGATTTTCAAGATGAGAGCTTCACTAAGCA 1508

QY 1693 GCAACACTATATATGTTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGA 1752
 DB 1509 GCAACAACTATATATGTTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGA 1568
 QY 1753 TATTTACGGGAAGGCTGTGCTGATGTGTGCTGGCCCGAAGCCTTACTGTGCTTGGGA 1812
 DB 1569 TATTTACGGGAAGGCTGTGCTGATGTGTGCTGGCCCGAAGCCTTACTGTGCTTGGGA 1628
 QY 1813 TGGTCTGATGTTCTGCTATATTTTCCCACTGCAAGAAGACGACAGACGACAGATAT 1872
 DB 1629 TGGTCTGATGTTCTGCTATATTTTCCCACTGCAAGAAGACGACAGACGACAGATAT 1688
 QY 1873 AAGAAATGAGAGCCCACTGACTGCTGTTCACTTACACTTACCACTGATATCAACATGAGCA 1932
 DB 1689 AAGAAATGAGAGCCCACTGACTGCTGTTCACTTACACTTACCACTGATATCAACATGAGCA 1748
 QY 1993 CAGCCCTGAGAGAAATCATCTATGTGTAGAGAAATGTAGCAATTTTGGAAATGAG 1992
 DB 1749 CAGCCCTGAGAGAAATCATCTATGTGTAGAGAAATGTAGCAATTTTGGAAATGAG 1808
 QY 1993 TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGGAAATGAGAGCGAA 2052
 DB 1809 TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGGAAATGAGAGCGAA 1868
 QY 2053 AGAAGAGATCAGAGTGTATGATCATATATCATGACAGACAGATCAAGGCTTCTGCTACGTAG 2112
 DB 1869 AGAAGAGATCAGAGTGTATGATCATATATCATGACAGACAGATCAAGGCTTCTGCTACGTAG 1928
 QY 2113 TCTACAAAGAAAGATTCAGGCAATTAACCTGCGCAATGCGGTGAAACATGGGTTATACA 2172
 DB 1929 TCTACAAAGAAAGATTCAGGCAATTAACCTGCGCAATGCGGTGAAACATGGGTTATACA 1988
 QY 2173 AACTCTTCTTAAGGTAAACCTGGAAGTATGACACAGAGCAATTTGAAAGAACTTCTCA 2232
 DB 1989 AACTCTTCTTAAGGTAAACCTGGAAGTATGACACAGAGCAATTTGAAAGAACTTCTCA 2048
 QY 2233 TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCATATACATGACACTAG 2292
 DB 2049 TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCATATACATGACACTAG 2108
 QY 2293 CCAGAGAGTCTGTGTCAGAGACTTATGACAGCTCATCAACCAACCCCAATCTCAACAGAT 2352
 DB 2109 CCAGAGAGTCTGTGTCAGAGACTTATGACAGCTCATCAACCAACCCCAATCTCAACAGAT 2168
 QY 2353 GGATGAGTCTGTGTCAGAGACTTGTGAAAGGAGCCGAAACCAAGTGGGCAAGGCCAGG 2412
 DB 2169 GGATGAGTCTGTGTCAGAGACTTGTGAAAGGAGCCGAAACCAAGTGGGCAAGGCCAGG 2228
 QY 2413 ACATACCCCAAGGAGACAGTAAACAAATGAGACCTTACAGAAATTAAGAAAGGTAGAA 2472
 DB 2229 ACATACCCCAAGGAGACAGTAAACAAATGAGACCTTACAGAAATTAAGAAAGGTAGAA 2288
 QY 2473 CAGAGAGACCAACGAATTTGAGAGGAGCCAGAGAGTGTCTGAGCTGCAATTAACCTTGA 2532
 DB 2289 CAGAGAGACCAACGAATTTGAGAGGAGCCAGAGAGTGTCTGAGCTGCAATTAACCTTGA 2348
 QY 2533 AACCTCAAAAGATGAGAACTTGCTAGACATTAAGTGAAGAAACAAATGCAATTAAT 2592
 DB 2349 AACCTCAAAAGATGAGAACTTGCTAGACATTAAGTGAAGAAACAAATGCAATTAAT 2408
 QY 2593 GAACCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
 DB 2409 GAACCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
 QY 2653 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2709
 DB 2469 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2525

RESULT 3
 US-60-172-373-8935
 ; Sequence 8935, Application US/60172373
 ; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonald
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Diep, Dinh
 ; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 ; FILE REFERENCE: GA-0006 P
 ; CURRENT FILING DATE: 1999-12-16
 ; NUMBER OF SEQ ID NOS: 25, 772
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 8935
 ; LENGTH: 2848
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 337522.1
 ; US-60-172-373-8935

Query Match 92.9%; Score 2517; DB 71; Length 2848;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACAGATGAGGCTGTGTTAACTAGATTTCTGTTCTTTCTGGGAGTATTAATTACAGC 252
 DB 9 CTGACAGATGAGGCTGTGTTAACTAGATTTCTGTTCTTTCTGGGAGTATTAATTACAGC 68
 QY 253 AAGAGCAAACTATAGAGATGAGAGAAACATGTGCGCAAGGCTGAATTTCTCAAGA 312
 DB 69 AAGAGCAAACTATAGAGATGAGAGAAACATGTGCGCAAGGCTGAATTTCTCAAGA 128
 QY 313 AATGTTGAATTCACAAATGATGATCACTTCAATGCTGTGGCCAAAGCTCCAGTTACA 372
 DB 129 AATGTTGAATTCACAAATGATGATCACTTCAATGCTGTGGCCAAAGCTCCAGTTACA 188
 QY 373 TACCTTCTTTTGGATGAGAGACGAGTACGCTGTATGTGAGCAAGATCAATATT 432
 DB 189 TACCTTCTTTTGGATGAGAGACGAGTACGCTGTATGTGAGCAAGATCAATATT 248
 QY 433 TTCAATGACCTGTGTTAATTAATCAAGATTTTCAAAATGTGTGCGCAATTTTAC 492
 DB 249 TTCAATGACCTGTGTTAATTAATCAAGATTTTCAAAATGTGTGCGCAATTTTAC 308
 QY 493 CAGAGAGATGAGAGCAAGTGGGCTGGAAGAAACATCTGGAAGAAATGTGCTAATTTCA 552
 DB 309 CAGAGAGATGAGAGCAAGTGGGCTGGAAGAAACATCTGGAAGAAATGTGCTAATTTCA 368
 QY 553 CAAGGTACTTAAGGATTAATTAATCAAGTCACTTGTACGCTGTGGAACGGGGCTTTTCA 612
 DB 369 CAAGGTACTTAAGGATTAATTAATCAAGTCACTTGTACGCTGTGGAACGGGGCTTTTCA 428
 QY 613 TCCAAATTTGACCTTACATTTGAAATTTGACATCACTTGAAGACAAATTTTAAAGCTGGA 672
 DB 429 TCCAAATTTGACCTTACATTTGAAATTTGACATCACTTGAAGACAAATTTTAAAGCTGGA 488
 QY 673 GAACCTCAATTTTGAAGACGGCGTGGGAAGAGTCCATATGACCCTTAAGCTGAGACG 732
 DB 489 GAACCTCAATTTTGAAGACGGCGTGGGAAGAGTCCATATGACCCTTAAGCTGAGACG 548
 QY 733 ATCCCTTTTAATAGATGAGAAATTAATCTGGAATGACGCTGATTTTATGGGGGAGA 792
 DB 549 ATCCCTTTTAATAGATGAGAAATTAATCTGGAATGACGCTGATTTTATGGGGGAGA 608
 QY 793 CTGTGCTATCTTCCGAATCTGTGGGACCAACCCCAATGAGACAGAGAGCATGATTC 852
 DB 609 CTGTGCTATCTTCCGAATCTGTGGGACCAACCCCAATGAGACAGAGAGCATGATTC 668
 QY 853 CAGGTGCTCAATGATCAAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGAATTC 912
 DB 669 CAGGTGCTCAATGATCAAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGAATTC 728
 QY 913 TGAAGATGACAAAGTATTAATTTTCTCCGTAAGAAATGCAATAGATGAGAGACATCTGG 972

Dp	729	TCAGAGTGCAGAAAGTAACTATTTTCTCCGTGAAAATGAAATGAGATGAGAGCACTCTGG	788
Qy	973	AAAAGCTACTCACGCTAGATAGCTCAGATATGCAAGATGACTTTGGAGGCGACAGAG	10322
Dp	789	AAAAGCTACTCACGCTAGAAATGAGTCAGATATGCAAGATGACTTTGGAGGCGACAGAG	848
Qy	1033	TCGTGGAAATTAATGAGCAACATTCCTCAAACTGCTCGATTTGCTCAGTGGCCAGGCTC	10922
Dp	849	TCGTGGAAATTAATGAGCAACATTCCTCAAACTGCTCGATTTGCTCAGTGGCCAGGCTC	908
Qy	1093	AAATGGCAATTGCACTCATTTTGTATGAACCTGCAAGATGTATTCCTAATGAACCTTTAAGA	11522
Dp	909	AAATGGCAATTGCACTCATTTTGTATGAACCTGCAAGATGTATTCCTAATGAACCTTTAAGA	968
Qy	1153	TCCTTAAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCAGATAACATTTTCAAGGATC	12122
Dp	969	TCCTTAAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCAGATAACATTTTCAAGGATC	10282
Qy	1213	AGCCGAGTGTATGTATGAGATGAGATGTGAGAGGGGTTCCTTGGTCCATATAGGCCA	12722
Dp	1029	AGCCGAGTGTATGTATGAGATGAGATGTGAGAGGGGTTCCTTGGTCCATATAGGCCA	10882
Qy	1273	CAGGAGTGGACCCCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCACGGCC	13322
Dp	1089	CAGGAGTGGACCCCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCACGGCC	11482
Qy	1333	AGGAACCTGTCCACAGAAAACATTTGGTGTGTGACTCTACAAAGACCTTCTGATGA	13922
Dp	1149	AGGAACCTGTCCACAGAAAACATTTGGTGTGTGACTCTACAAAGACCTTCTGATGA	12082
Qy	1393	TGTTTAACTCTTGAAGAAGTCAACGACATGACATCAATCAGTGTTCCTATGACAA	14522
Dp	1209	TGTTTAACTCTTGAAGAAGTCAACGACATGACATCAATCAGTGTTCCTATGACAA	12682
Qy	1453	TCGCCCAATAGTATCAAAACGAGATGTAATTAATCAATTTACAACATTTGTCGTAGACG	15122
Dp	1259	TCGCCCAATAGTATCAAAACGAGATGTAATTAATCAATTTACAACATTTGTCGTAGACG	13282
Qy	1513	AGTGGATGCAGAAAGTGCAGATATGATGTATATGTTTATCCGAAACAGATGTTGGACCGT	15722
Dp	1339	AGTGGATGCAGAAAGTGCAGATATGATGTATATGTTTATCCGAAACAGATGTTGGACCGT	13882
Qy	1573	TCTTAAATAGTTTCAATTCCTAAGAGACCTTGTATATGATTAAGAAGGTTCTGCTGGA	16322
Dp	1389	TCTTAAATAGTTTCAATTCCTAAGAGACCTTGTATATGATTAAGAAGGTTCTGCTGGA	14482
Qy	1633	AGAAATGACAGTTTTTCGGGAAACGACATGCTATTTCAAGCAATGAGCTTTCCACTAAGCA	16922
Dp	1449	AGAAATGACAGTTTTTCGGGAAACGACATGCTATTTCAAGCAATGAGCTTTCCACTAAGCA	15082
Qy	1693	GCAACACTATATATATGTTGTTCAACGGCTGGGGTGTCCACAGCTCCCTTTTACAACGGTGTGA	17522
Dp	1509	GCAACACTATATATATGTTGTTCAACGGCTGGGGTGTCCACAGCTCCCTTTTACAACGGTGTGA	15682
Qy	1753	TATTTAAGGGAAGAGGTGTGCTGAGTGTTCGCTGCGCCGAGACCTTATCTGTGCTGGGA	18122
Dp	1569	TATTTAAGGGAAGAGGTGTGCTGAGTGTTCGCTGCGCCGAGACCTTATCTGTGCTGGGA	16282
Qy	1813	TGTTTCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGACAAGACGACAAGATAT	18722
Dp	1629	TGTTTCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGACAAGACGACAAGATAT	16882
Qy	1873	AAGAAATGAGAGCCCACTGACTCATGTGTTCAGACTTACACCATGATTAATCAACATGGCCA	19322
Dp	1689	AAGAAATGAGAGCCCACTGACTCATGTGTTCAGACTTACACCATGATTAATCAACATGGCCA	17482
Qy	1933	CAGCCCTGAAGAGATCATCTATGTGTGTAGAGATATGTACACATTTTGGAAATGCAG	19922
Dp	1749	CAGCCCTGAAGAGATCATCTATGTGTGTAGAGATATGTAGCAATTTTGGAAATGCAG	18082
Qy	1993	TCCGAAGTCGAGAGAGGCTGTGCTATTTGGCAATTCAGAGGGGAAATGAAAGGCCAAA	20522
Dp	1809	TCCGAAGTCGAGAGAGGCTGTGCTATTTGGCAATTCAGAGGGGCAAAATGAAAGGCCAAA	18682

QY	2053	AGAAAGATCAAGTGGATATCATATCATACAGGACAATCAAGGCTTCGTCTACGAG	2113
Db	1869	AGAAAGATCAAGTGGATATCATATCATACAGGACAATCAAGGCTTCGTCTACGAG	1922
QY	2113	TCTCAACAAGAGATTCAGGCAATTACCTCTCGCATCGGTGGAACTGGGTTTATCA	2172
Db	1929	TCTCAACAAGAGATTCAGGCAATTACCTCTCGCATCGGTGGAACTGGGTTTATCA	1988
QY	2173	AACCTCTCTTAAGGTAAACCTCGGAAGTCATGTGACAAGAGATTTGGAGAACTTC	2233
Db	1989	AACCTCTCTTAAGGTAAACCTCGGAAGTCATGTGACAAGAGATTTGGAGAACTTC	2044
QY	2233	TAAAGATGATGAGAGATGCGCTCTAAGACCAAGAAATGTCCAATACATGACACT	2293
Db	2049	TAAAGATGATGAGAGATGCGCTCTAAGACCAAGAAATGTCCAATACATGACACT	2108
QY	2293	CCAGAAGTCTGCTACAGAGACTTCATGACGCTCATCAACCCCAATCTTCACACAT	2353
Db	2109	CCAGAAGTCTGCTACAGAGACTTCATGACGCTCATCAACCCCAATCTTCACACAT	2168
QY	2353	GGATGAGTTCGTGTAACAAGTTTGGAAAGGGAACCGAAACAAAGTGGGCAAGGCAAG	2413
Db	2169	GGATGAGTTCGTGTAACAAGTTTGGAAAGGGAACCGAAACAAAGTGGGCAAGGCAAG	2222
QY	2413	ACATATCCCGAAGGAACAGTAAACAATGGAGACCTTACAGAAATATGAAGAAGTAGAA	2472
Db	2229	ACATATCCCGAAGGAACAGTAAACAATGGAGACCTTACAGAAATATGAAGAAGTAGAA	2288
QY	2473	CAGAGAGACCCACGAATTTGAGAGGGAACCCAGAGATGTCTGAGTGCATTAACCTCTGA	2533
Db	2289	CAGAGAGACCCACGAATTTGAGAGGGAACCCAGAGATGTCTGAGTGCATTAACCTCTGA	2348
QY	2533	AACCTCAAAACAAGTAGAACTTGCTGCTGACATATACATTAACGTGAAATATACAT	2593
Db	2349	AACCTCAAAACAAGTAGAACTTGCTGCTGACATATACATTAACGTGAAATATACAT	2408
QY	2593	GAACTTTTTCATGSCATTAATGAGATGTTTACAAATGGTGGAAATTCAGCTGAGTTCCA	2653
Db	2409	GAACTTTTTCATGSCATTAATGAGATGTTTACAAATGGTGGAAATTCAGCTGAGTTCCA	2468
QY	2653	CCAAATTATTAATTTAAATCCAGATTAACCTTCCAAATAGGCTTTTTCCTTAATACC	2709
Db	2469	CCAAATTATTAATTTAAATCCAGATTAACCTTCCAAATAGGCTTTTTCCTTAATACC	2525

RESULT 4
US-60-323-784-83

```

GENERAL INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 P
CURRENT APPLICATION NUMBER: US/60/323,784
CURRENT FILING DATE: 2001-09-19
NUMBER OF SEQ. ID NOS.: 186
SOFTWARE: PERL Program
SEQ ID NO 83
LENGTH: 2848
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1930967CBA1
US-60-323-784-83

```

Query Match	92.9%	Score 2517	DB 87	Length 2848
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2517	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 193 CTGCGCATGGGCTGGTACTAGATGTTGTCCTTTCTGGGAGATTACTTAACAC 252
Db 9 CTGCGCATGGGCTGGTACTAGATGTTGTCCTTTCTGGGAGATTACTTAACAC 68
QY 253 AAGACCAACTATCAGAAATGGAGAACAAATGGCCAGGCTGAATTAATCTTACAAAG 312
Db 69 AAGACCAACTATCAGAAATGGAGAACAAATGGCCAGGCTGAATTAATCTTACAAAG 128
QY 313 AATGTGGATCCAACTATGATGATCACTTCAATGGCTGGCCAAACAGCTCAGATTACA 372
Db 129 AATGTGGATCCAACTATGATGATCACTTCAATGGCTGGCCAAACAGCTCAGATTACA 188
QY 373 TACCTTCTTTGGATGAGAGACGAGATGCTGATGTTGGAGCAAAAGATCACAATTT 432
Db 189 TACCTTCTTTGGATGAGAGACGAGATGCTGATGTTGGAGCAAAAGATCACAATTT 248
QY 433 TTCAATTCACCTGGTAAATATCAAGATTTTCAAAAGATGTTGGCCAGATCTTAAC 492
Db 249 TTCAATTCACCTGGTAAATATCAAGATTTTCAAAAGATGTTGGCCAGATCTTAAC 308
QY 493 CAGAGAGATGATCAAGTGGGCTGAAAAGACATCTGAAAAGATGTGCTAATTTCAAT 552
Db 309 CAGAGAGATGATCAAGTGGGCTGAAAAGACATCTGAAAAGATGTGCTAATTTCAAT 368
QY 553 CAGGATCTTAAGCATATATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAGGATCTTAAGCATATATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 428
QY 613 TTCAATTCACCTGGTAAATATCAAGATTTTCAAAAGATGTTGGCCAGATCTTAAC 672
Db 429 TTCAATTCACCTGGTAAATATCAAGATTTTCAAAAGATGTTGGCCAGATCTTAAC 488
QY 673 GAACTCACATTTTGAAGACGGCCGTGGAGAGATCATATGACCTTAAGCTGTGACAGC 732
Db 489 GAACTCACATTTTGAAGACGGCCGTGGAGAGATCATATGACCTTAAGCTGTGACAGC 548
QY 733 ATCCCTTTTAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGCGAGA 792
Db 549 ATCCCTTTTAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGCGAGA 608
QY 793 CTGTCCTATCTTCCGAACCTCTGGGACACCAACCATTCAGAGACAGACATGATTC 852
Db 609 CTGTCCTATCTTCCGAACCTCTGGGACACCAACCATTCAGAGACAGACATGATTC 668
QY 853 CAGGTGCTCATATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGATGACATCC 912
Db 669 CAGGTGCTCATATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGATGACATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTCTTCGTTGAAAATGCAATGATGAGAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTCTTCGTTGAAAATGCAATGATGAGAACACTCTGG 788
QY 973 AAAAGCTATCTACGCTAGATATGCTCAGATATGCAAGATGACTTTGGAGGCGACAGAG 1032
Db 789 AAAAGCTATCTACGCTAGATATGCTCAGATATGCAAGATGACTTTGGAGGCGACAGAG 848
QY 1033 TCTGTGATTAATGAGCAACATCTCTCAAGCTGCTGATTTGCTCAGTCCAGGCTCC 1092
Db 849 TCTGTGATTAATGAGCAACATCTCTCAAGCTGCTGATTTGCTCAGTCCAGGCTCC 908
QY 1093 AAATGGCATGACACTCATTTTGGATGAACCTGAGAGATGATCTTAATGAACCTTTAAAG 1152
Db 909 AAATGGCATGACACTCATTTTGGATGAACCTGAGAGATGATCTTAATGAACCTTTAAAG 968
QY 1153 TCTTAAAAATCAGATGATATGAGATGTTTGAAGCTTCAGATTAACATTTTCAAGGATC 1212
Db 969 TCTTAAAAATCAGATGATATGAGATGTTTGAAGCTTCAGATTAACATTTTCAAGGATC 1028
QY 1213 AGCCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db 1029 AGCCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY 1273 CAGGATGAGACCAACTATCAATGAGGCTGCTTATCAAGAGAGATGCCCTATCAACGCGC 1332

Db 1089 CAGGATGAGACCAACTATCAATGAGGCTGCTTATCAAGAGAGATGCCCTATCAACGCGC 1148
QY 1333 AGGAATGCTGCCAGCAAAACATTTGGTGGTTTGAAGCTTCAAGAGACCTTCTGATGA 1392
Db 1149 AGGAATGCTGCCAGCAAAACATTTGGTGGTTTGAAGCTTCAAGAGACCTTCTGATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGATCATCAGCCATGATCAATCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTGCAAGAGATCATCAGCCATGATCAATCAGTGTTCCTATGAACAA 1268
QY 1453 TCGCCCAATGATGATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1512
Db 1269 TCGCCCAATGATGATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1328
QY 1513 AGTGGATGACAGAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db 1329 AGTGGATGACAGAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
QY 1573 TCTTAAAGTACTTCAATCTTAAAGAGACTTGGTATGATTAAGAGAGTTCGTGCTGA 1632
Db 1389 TCTTAAAGTACTTCAATCTTAAAGAGACTTGGTATGATTAAGAGAGTTCGTGCTGA 1448
QY 1633 AGAATGACAGTTCCTTGGAGACCTGCTATTCAGCAATGAGACTTTCATTAAGCA 1692
Db 1449 AGAATGACAGTTCCTTGGAGACCTGCTATTCAGCAATGAGACTTTCATTAAGCA 1508
QY 1693 GCAACACTATATATTTGGTTCAGAGGCTGGGCTGGCCAGCTCCCTTAACAGCGGTGGA 1752
Db 1509 GCAACACTATATATTTGGTTCAGAGGCTGGGCTGGCCAGCTCCCTTAACAGCGGTGGA 1568
QY 1753 TATTTAAGGAGAAACGTCGTGAGATGTCCTGCGCCAGAGACCTTACTGTGCTGGA 1812
Db 1569 TATTTAAGGAGAAACGTCGTGAGATGTCCTGCGCCAGAGACCTTACTGTGCTGGA 1628
QY 1813 TGTTCCTGATGTCCTGCTATTTTCCACTGCAAGAGACGCAACAGACAGATAT 1872
Db 1629 TGTTCCTGATGTCCTGCTATTTTCCACTGCAAGAGACGCAACAGACAGATAT 1688
QY 1873 AAGAAATGAGAGACCACTGACTCATCTGTTGCAAGCTTACCATGATATACCAATGGCCA 1932
Db 1689 AAGAAATGAGAGACCACTGACTCATCTGTTGCAAGCTTACCATGATATACCAATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAGATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
Db 1749 CAGCCCTGAAGAGAGATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
QY 1993 TCCGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTCGAGGCGGAATGAAGCGAAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTCGAGGCGGAATGAAGCGAAA 1868
QY 2053 AGAAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
Db 1869 AGAAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
QY 2113 TCTTCAACAGAGAGATTCAGGCAATTAATCTGCTGCAAGCGGTGGAACATGGGTTTATCA 2172
Db 1929 TCTTCAACAGAGAGATTCAGGCAATTAATCTGCTGCAAGCGGTGGAACATGGGTTTATCA 1988
QY 2173 AACTCTTTCTTAAGTAAACCTGGAAGTCAATTTGACAGAGACTTTGGAAGAACTTCTTCA 2232
Db 1989 AACTCTTTCTTAAGTAAACCTGGAAGTCAATTTGACAGAGACTTTGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
Db 2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108
QY 2293 CCAAGAGTCTGATGACAGAGACTTCAAGAGCTGATCAACCAACCCATCTCAACAGAT 2352
Db 2109 CCAAGAGTCTGATGACAGAGACTTCAAGAGCTGATCAACCAACCCATCTCAACAGAT 2168
QY 2353 GGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2412

Db 2169 GGATGATTCTGTGAACAAATTGGAAAAAGGACCGAAAAACAAGTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCGAGGAGCAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAAA 2472
Db 2229 ACATACCCGAGGAGCAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAAA 2288
QY 2473 CAGGAGAGCCAGCAATTTGAGAGGCAACCAGAGTGTCTGAGCTGCAATTACTCTAGA 2532
Db 2289 CAGGAGAGCCAGCAATTTGAGAGGCAACCAGAGTGTCTGAGCTGCAATTACTCTAGA 2348
QY 2533 AACCTTAACCAAGTAGAACTTGCCTGACATTAATCTGAGAAAAACAATGCAATTAACAT 2592
Db 2349 AACCTTAACCAAGTAGAACTTGCCTGACATTAATCTGAGAAAAACAATGCAATTAACAT 2408
QY 2593 GAACCTTTTCAAGGATTAATGATGATGATTTTACAAATGATGAGAAATTCAGCTGATTCGA 2652
Db 2409 GAACCTTTTCAAGGATTAATGATGATGATTTTACAAATGATGAGAAATTCAGCTGATTCGA 2468
QY 2653 CCAATTATTAATTAATTCATGAGTAACCTTCTTAATAGGCTTTTCTTATATACC 2709
Db 2469 CCAATTATTAATTAATTCATGAGTAACCTTCTTAATAGGCTTTTCTTATATACC 2525

RESULT 5
PCT-US02-07826-283

/ Sequence 283, Application PC/TUS0207826

/ GENERAL INFORMATION:

/ APPLICANT: Millennium Pharmaceuticals, Inc. et al.

/ TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,

/ FILE REFERENCE: MRI-030PC

/ CURRENT FILING DATE: 2002-03-14

/ PRIOR FILING DATE: 2001-03-14

/ PRIOR APPLICATION NUMBER: 60/276,025

/ PRIOR FILING DATE: 2001-03-14

/ PRIOR APPLICATION NUMBER: 60/325,149

/ PRIOR FILING DATE: 2001-09-27

/ PRIOR APPLICATION NUMBER: 60/276,026

/ PRIOR FILING DATE: 2001-03-14

/ PRIOR APPLICATION NUMBER: 60/324,967

/ PRIOR FILING DATE: 2001/09/26

/ PRIOR APPLICATION NUMBER: 60/311,732

/ PRIOR FILING DATE: 2001-08-10

/ PRIOR APPLICATION NUMBER: 60/325,102

/ PRIOR FILING DATE: 2001-09-26

/ PRIOR APPLICATION NUMBER: 60/323,580

/ PRIOR FILING DATE: 2001-09-19

/ NUMBER OF SEQ ID NOS: 363

/ SOFTWARE: FastSeq for windows Version 4.0

/ SEQ ID NO 283

/ TYPE: DNA

/ LENGTH: 2530

/ ORGANISM: Homo sapiens

PCT-US02-07826-283

Query Match 92.6%; Score 2508; DB 1; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACGATGAGGCTGTACTAGATGATGCTCTTTCTTTCTGGGAGATTAATTACAGC 252
Db 9 CTGACGATGAGGCTGTACTAGATGATGCTCTTTCTTTCTGGGAGATTAATTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAGAAACAATGGCCAAAGCTGAATTAATCCACAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAGAAACAATGGCCAAAGCTGAATTAATCCACAAAGA 128
QY 313 AATGTTGAATCCAAATGATGATCACTTCAATGAGCTTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTTGAATCCAAATGATGATCACTTCAATGAGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTTGGATGAGAAAGGAGTGGTGTATGTTGGAGCAAGATCAATATT 432

Db 189 TACCTTCTTTTGGATGAGAAAGGAGTGGTGTATGTTGGAGCAAGATCAATATT 248
QY 433 TTCAATTCGACCTGGTTAATATCAAGATTTTCAAAAGATGTTGTGGCCAGTATCTTAAC 492
Db 249 TTCAATTCGACCTGGTTAATATCAAGATTTTCAAAAGATGTTGTGGCCAGTATCTTAAC 308
QY 493 CAGAAAGATGATGCAAGTGGGCTGAAAGACATCTGAAAGATGCTAATTTTCAAT 552
Db 309 CAGAAAGATGATGCAAGTGGGCTGAAAGACATCTGAAAGATGCTAATTTTCAAT 368
QY 553 CAGGATCTTAAGGATTAATATCAAGATCTTGTAGCCCTGTGGAAGGGGGCTTTTCA 612
Db 369 CAGGATCTTAAGGATTAATATCAAGATCTTGTAGCCCTGTGGAAGGGGGCTTTTCA 428
QY 613 TCCAAATTCGACCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
Db 429 TCCAAATTCGACCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 488
QY 673 GAACCTCAATTTTGAAGAAAGGCGGTGGGAGAGTGCATATGACCTTAAGCTGACAGC 732
Db 489 GAACCTCAATTTTGAAGAAAGGCGGTGGGAGAGTGCATATGACCTTAAGCTGACAGC 548
QY 733 ATCCCTTTTAAATGATGAGAAATTAATCTGGAATCTGAGTGAATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTTAAATGATGAGAAATTAATCTGGAATCTGAGTGAATTTTATGGGGCGAGA 608
QY 793 CTTGCTATCTTCCGAACTCTTGGGACCAACCAATCAGAGCAGAGCATGATTC 852
Db 609 CTTGCTATCTTCCGAACTCTTGGGACCAACCAATCAGAGCAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCAATCTCAGAGAGTGAATTC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCAATCTCAGAGAGTGAATTC 728
QY 913 TGAAGATGCAAAAGATTAATCTTTTCTCCGTAAGAAATGCAATTAATGAGAAATCTCGG 972
Db 729 TGAAGATGCAAAAGATTAATCTTTTCTCCGTAAGAAATGCAATTAATGAGAAATCTCGG 788
QY 973 AAAAGCTACTACGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
Db 789 AAAAGCTACTACGCTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
QY 1033 TCTGTGATTAATTAATGAGCAATTCCTGAAGCTGCTGATTTGCTCAGTGCAGCTCC 1092
Db 849 TCTGTGATTAATTAATGAGCAATTCCTGAAGCTGCTGATTTGCTCAGTGCAGCTCC 908
QY 1093 AATGCGATGACACTCTTTTGTGATGAATCTGAGAGATGATTTCTTAATGAATTTAAAGA 1152
Db 909 AATGCGATGACACTCTTTTGTGATGAATCTGAGAGATGATTTCTTAATGAATTTAAAGA 968
QY 1153 TCTTAATAATCAAGTGTATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Db 969 TCTTAATAATCAAGTGTATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
QY 1213 AGCGGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db 1029 AGCGGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY 1273 CAGGATGAGACCAACTATCAATGAGTGCCTTATCAAGAGAGATGCCCTATCCAGGCC 1332
Db 1089 CAGGATGAGACCAACTATCAATGAGTGCCTTATCAAGAGAGATGCCCTATCCAGGCC 1148
QY 1333 AGGAATCTGTCCAGCAAAATTTGGTGGTTTGAATCTCAAAAGAGCTTCTCGATGA 1392
Db 1149 AGGAATCTGTCCAGCAAAATTTGGTGGTTTGAATCTCAAAAGAGCTTCTCGATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCAAGTGTCTTCTATGAACA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCAAGTGTCTTCTATGAACA 1268
QY 1453 TCGCCCAATAGTATCAAAAGGATGTAAATTTATCAATTTATCAAAATTTGTCGTAGACG 1512
Db 1269 TCGCCCAATAGTATCAAAAGGATGTAAATTTATCAATTTATCAAAATTTGTCGTAGACG 1328


```
APPLICANT: Karen GIATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SPO ID NO 283
LENGTH: 2530
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-283

Query Match      92.6%; Score 2508; DB 45; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 CTGACGATGGGCTGTTAATCTAGAGTTGTCTCTTTCTGGGGAGTATCTTACAGC 252
DB 9 CTGCGACATGGGCTGTTAATCTAGAGTTGTCTCTTTCTGGGGAGTATCTTACAGC 68

QY 253 AAGACCAACTATCAGATGGAGAGAACATGTGCGCAAGCTGAAATATCTTACAAAGA 312
DB 69 AAGACCAACTATCAGATGGAGAGAACATGTGCGCAAGCTGAAATATCTTACAAAGA 128

QY 313 AATGTTGGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTTCAATATCA 372
DB 129 AATGTTGGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTTCAATATCA 188

QY 373 TACCTTCCCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
DB 189 TACCTTCCCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248

QY 433 TTCAATTCGACCTGTTAATATCAAGAGATTTTCAAAAGATTTGTGGCCAGTATCTTAC 492
DB 249 TTCAATTCGACCTGTTAATATCAAGAGATTTTCAAAAGATTTGTGGCCAGTATCTTAC 308

QY 493 CAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
DB 309 CAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368

QY 553 CAAAGGACTAGAGATATATCAGACTGATCTGTAAGCGCTGGAAGCGGGGCTTTTCA 612
DB 369 CAAAGGACTAGAGATATATCAGACTGATCTGTAAGCGCTGGAAGCGGGGCTTTTCA 428

QY 613 TCCAAATTTGACCTATCAATTAATGAGATCATCTGAGAGCAATATTTTAACTGGA 672
DB 429 TCCAAATTTGACCTATCAATTAATGAGATCATCTGAGAGCAATATTTTAACTGGA 488

QY 673 GAATCTACATTTTGAAGAGCGCGTGGAGAGAGTCCATATGACCTTAAGCTGACAGC 732
DB 489 GAATCTACATTTTGAAGAGCGCGTGGAGAGAGTCCATATGACCTTAAGCTGACAGC 548

QY 733 ATCCCTTTAATAGAGAGATTAATCTGGAAGCTGAGAGCTGATTTTAAAGGAGAGA 792
DB 549 ATCCCTTTAATAGAGAGATTAATCTGGAAGCTGAGAGCTGATTTTAAAGGAGAGA 608

QY 793 CTTTGCTATCTTCGAACTCTTGAGGACCAACCAATCAGAGAGAGAGAGAGATTC 852
```

```
DB 609 CTTTGCTATCTTCGAACTCTTGAGGACCAACCAATCAGAGAGAGAGAGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTAAGGCCACCTCATCTCAGAGATGACAAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTAAGGCCACCTCATCTCAGAGATGACAAATCC 728
QY 913 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAAATGCAATGAGAGAACTCTGG 972
DB 729 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAAATGCAATGAGAGAACTCTGG 788
QY 973 AAAAGCTACTACAGCTAGAAATAGGTCAATATGCAAGATGACTTTGGAGGACAGAG 1032
DB 789 AAAAGCTACTACAGCTAGAAATAGGTCAATATGCAAGATGACTTTGGAGGACAGAG 848
QY 1033 TCTGATGATTAATGAGACAACTTCCTAAAGCTGCTGATTTGCTAGGCGCAGAGTCC 1092
DB 849 TCTGATGATTAATGAGACAACTTCCTAAAGCTGCTGATTTGCTAGGCGCAGAGTCC 908
QY 1093 AAATGGCATGACACTATTTTGTATGAACTGACAGAGATGATTCCTAATGAACTTTAAGA 1152
DB 909 AAATGGCATGACACTATTTTGTATGAACTGACAGAGATGATTCCTAATGAACTTTAAGA 968
QY 1153 TCTTAAATATCCAGTTGATATGAGAGTGTTAACACTTCCAGTAACTTTTCAAGGAGATC 1212
DB 969 TCTTAAATATCCAGTTGATATGAGAGTGTTAACACTTCCAGTAACTTTTCAAGGAGATC 1028
QY 1213 AGCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
DB 1029 AGCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY 1273 CAGGATGAGACCAACTATCATGAGGCTTATCAAGAGAGATCCCTATCCAGGCGC 1332
DB 1089 CAGGATGAGACCAACTATCATGAGGCTTATCAAGAGAGATCCCTATCCAGGCGC 1148
QY 1333 AGGAACTTGTCCAGCAAAACATTTGTGTTTGACTTCAAAAGAGCTTCTGTATGA 1392
DB 1149 AGGAACTTGTCCAGCAAAACATTTGTGTTTGACTTCAAAAGAGCTTCTGTATGA 1208
QY 1393 TGTATTAACCTTTTGCAGAAATGATCCAGGATGATGATGATGATGATGATGATGATGAT 1452
DB 1209 TGTATTAACCTTTTGCAGAAATGATCCAGGATGATGATGATGATGATGATGATGATGAT 1268
QY 1453 TCGCCCAATGATGATCAAAACGAGATGATTAATCAATTTTCACAAAATTTGTGTAAG 1512
DB 1269 TCGCCCAATGATGATCAAAACGAGATGATTAATCAATTTTCACAAAATTTGTGTAAG 1328
QY 1513 AGTGAATGACAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
DB 1329 AGTGAATGACAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
QY 1573 TCTTAAGTATTTCAATTTCTTAAGAGAGACTTGGTATGATTAAGAGAGTTCGCTGGA 1632
DB 1389 TCTTAAGTATTTCAATTTCTTAAGAGAGACTTGGTATGATTAAGAGAGTTCGCTGGA 1448
QY 1633 AGAATGACAGTTTTCGGAAGACGAGCTGATTTTCAAGATGAGAGCTTCCATTAAGCA 1692
DB 1449 AGAATGACAGTTTTCGGAAGACGAGCTGATTTTCAAGATGAGAGCTTCCATTAAGCA 1508
QY 1693 GCAACAACTATATATTTGTTGAACGCTGGGATTTGCCAGCTCCCTTTTCAACCGGTGGA 1752
DB 1509 GCAACAACTATATATTTGTTGAACGCTGGGATTTGCCAGCTCCCTTTTCAACCGGTGGA 1568
QY 1753 TATTTAAGGAGAAAGCGTGTGAGTGTGCTGCTGCGCCGAGACCTTACTGTGCTGGGA 1812
DB 1569 TATTTAAGGAGAAAGCGTGTGAGTGTGCTGCTGCGCCGAGACCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCTGACATGTTCTGCTATTTTCCACATGCAAGAGAGAGCAAGAGAGATAT 1872
DB 1629 TGGTTCTGACATGTTCTGCTATTTTCCACATGCAAGAGAGAGCAAGAGAGATAT 1688
QY 1873 AAGAAATGAGACCACTGACTGACTGTTGAGACTTACCATGATATATCAATGAGCA 1932
```

Db 1689 AAGAAATGAGACCCACTGACCTGTTGACGTTACACCATGATATATCAATGAGCCA 1748
 QY 1933 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGAAATAGAGACATTTTGGAAATGAG 1992
 Db 1749 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGAAATAGAGACATTTTGGAAATGAG 1808
 QY 1993 TCCGAAGTGCAGAGAGCGCTGCTCTATTGCGCAATTCAGAGGCGAAATGAGAGCGAAA 2052
 Db 1809 TCCGAAGTGCAGAGAGCGCTGCTCTATTGCGCAATTCAGAGGCGAAATGAGAGCGAAA 1868
 QY 2053 AGAAGAGATCAGAGTGTATATATATCATCAGAGACATGAGGCTCTGCTAGCTGAG 2112
 Db 1869 AGAAGAGATCAGAGTGTATATATATCATCAGAGACATGAGGCTCTGCTAGCTGAG 1928
 QY 2113 TCTACACAGAGAGATTCAGAGCAATTAACCTGCGCATGCGGTGAAACATGGCTTCAATCA 2172
 Db 1929 TCTACACAGAGAGATTCAGAGCAATTAACCTGCGCATGCGGTGAAACATGGCTTCAATCA 1988
 QY 2173 AACTCTTCTTAAGGTAACTCTGAGAGTCAATTAACAGAGCAATTTGAGAGAACTTCTTCA 2232
 Db 1989 AACTCTTCTTAAGGTAACTCTGAGAGTCAATTAACAGAGCAATTTGAGAGAACTTCTTCA 2048
 QY 2233 TAAAGATGATGAGAGATGAGTCTTAAGACCAAGAAATGTCCAATAGATGACACTGAG 2292
 Db 2049 TAAAGATGATGAGAGATGAGTCTTAAGACCAAGAAATGTCCAATAGATGACACTGAG 2108
 QY 2293 CCAGAGAGTCTGTGTAGAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGAT 2352
 Db 2109 CCAGAGAGTCTGTGTAGAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGAT 2168
 QY 2353 GAGTAGTCTGTGAACAAAGTTGAGAAAGGAGCCGAAACAAAGTGTGAGAAAGGCGCAGG 2412
 Db 2169 GAGTAGTCTGTGAACAAAGTTGAGAAAGGAGCCGAAACAAAGTGTGAGAAAGGCGCAGG 2228
 QY 2413 AACTATCCCAAGAGAAAGATTAACAAATGAGAGCACTTAACAGAAATTAAGAAAGGTGAAA 2472
 Db 2229 AACTATCCCAAGAGAAAGATTAACAAATGAGAGCACTTAACAGAAATTAAGAAAGGTGAAA 2288
 QY 2473 CAGAGAGACCAAGATTTGAGAGGCAACCCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2532
 Db 2289 CAGAGAGACCAAGATTTGAGAGGCAACCCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2348
 QY 2533 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATTAACCTGAGAAAAACAAATGCAATATATCAT 2592
 Db 2349 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATTAACCTGAGAAAAACAAATGCAATATATCAT 2408
 QY 2593 GAACCTTTTTCATGAGATTAATGAGATGTTTACAAATGAGGAAATTCAGCTGAGTTCCA 2652
 Db 2409 GAACCTTTTTCATGAGATTAATGAGATGTTTACAAATGAGGAAATTCAGCTGAGTTCCA 2468
 QY 2653 CCAATATTAATTAATCAATGAGTAACTTCTTAATAGGCTTTTCTT 2700
 Db 2469 CCAATATTAATTAATCAATGAGTAACTTCTTAATAGGCTTTTCTT 2516

RESULT 8
 US-10-262-538-9
 ; Sequence 9, Application US/10262538

; GENERAL INFORMATION:
 ; APPLICANT: Alitalo et al
 ; TITLE OF INVENTION: NEUROPELIN/VBGF-C/VBGF-3 MATERIALS AND METHODS
 ; FILE REFERENCE: 28967/37564
 ; CURRENT APPLICATION NUMBER: US/10/262,538
 ; CURRENT FILING DATE: 2002-09-30
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 2530
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (16)..(2331)

US-10-262-538-9
 Query Match 92.6%; Score 2508; DB 49; Length 2530;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACAGATGGGCTGGTTACTAGGATGCTGTCTTTTCTGGGGAGTATTAATTACAGC 252
 Db 9 CTGACAGATGGGCTGGTTACTAGGATGCTGTCTTTTCTGGGGAGTATTAATTACAGC 68
 QY 253 AAGACAAACCTATCAGAAATGGGAAAGACAAATGTCAGAGGCGTGAATTAATCTACAAAGA 312
 Db 69 AAGACAAACCTATCAGAAATGGGAAAGACAAATGTCAGAGGCGTGAATTAATCTACAAAGA 128
 QY 313 AATGTTGATCCAAACAAATGATATCACTTCAATGCTGTGGCAACAGCTCCAGTTATCA 372
 Db 129 AATGTTGATCCAAACAAATGATATCACTTCAATGCTGTGGCAACAGCTCCAGTTATCA 188
 QY 373 TACCTTCTTTTGGATGAGGAAACGAGTAGGCTGTATGTGGAGCAAGAGATCAATATT 432
 Db 189 TACCTTCTTTTGGATGAGGAAACGAGTAGGCTGTATGTGGAGCAAGAGATCAATATT 248
 QY 433 TTCATTGACCTGGTTAATATCAAGATTTTCAAAAGATGTGTGCGCATCTTACAC 492
 Db 249 TTCATTGACCTGGTTAATATCAAGATTTTCAAAAGATGTGTGCGCATCTTACAC 308
 QY 493 CAGAGAGATGATCAAGTGGGCTGAGAAAGACATCTGTAAGAAAGATGTGTAATTTGAT 552
 Db 309 CAGAGAGATGATCAAGTGGGCTGAGAAAGACATCTGTAAGAAAGATGTGTAATTTGAT 368
 QY 553 CAGGTAATTAAGGATTAATATCAAGTCACTTGTAGCCTGTGTGAAACGGGGGCTTTTCA 612
 Db 369 CAGGTAATTAAGGATTAATATCAAGTCACTTGTAGCCTGTGTGAAACGGGGGCTTTTCA 428
 QY 613 TCCAAATTTGACCTTCAATTGAAATGGAATCATTACCTGAGACATATTTTAAAGCTGGA 672
 Db 429 TCCAAATTTGACCTTCAATTGAAATGGAATCATTACCTGAGACATATTTTAAAGCTGGA 488
 QY 673 GAACCTCACTTTTGAAGACGCGCTGGAGAGAGTGCATATGACCTTAAGCTGCTGACAGC 732
 Db 489 GAACCTCACTTTTGAAGACGCGCTGGAGAGAGTGCATATGACCTTAAGCTGCTGACAGC 548
 QY 733 ATCCCTTTTAATAGATGAGAAATTAATCTGGAACCTGACGTGATTTTATGGGCGAGA 792
 Db 549 ATCCCTTTTAATAGATGAGAAATTAATCTGGAACCTGACGTGATTTTATGGGCGAGA 608
 QY 793 CTTCCTATCTTCCCAACTCTTGGGACACCAACCAATACAGACAGACATGATTC 852
 Db 609 CTTCCTATCTTCCCAACTCTTGGGACACCAACCAATACAGACAGACATGATTC 668
 QY 853 CAGTGGCTCAATGATCCAAAGTTCAATAGTCCACACTCATCTCAGAGAGTGAACATCC 912
 Db 669 CAGTGGCTCAATGATCCAAAGTTCAATAGTCCACACTCATCTCAGAGAGTGAACATCC 728
 QY 913 TGAAGATGACAAAGTATACCTTTTCTCGTGAAGAAATGCAATAGATGAGAACTCTGG 972
 Db 729 TGAAGATGACAAAGTATACCTTTTCTCGTGAAGAAATGCAATAGATGAGAACTCTGG 788
 QY 973 AAAAGCTACACGCTTAATTAATGATGATGATGCAAGAAATGCAATTTGAGGGGACAGAG 1032
 Db 789 AAAAGCTACACGCTTAATTAATGATGATGATGCAAGAAATGCAATTTGAGGGGACAGAG 848
 QY 1033 TCTGTGATTAATAGACAAATTCCTCAAGAGTGTGTGATTTGCTCAGTGCAGGTC 1092
 Db 849 TCTGTGATTAATAGACAAATTCCTCAAGAGTGTGTGATTTGCTCAGTGCAGGTC 908
 QY 1093 AAATGGCATATGACATCTATTTGATGAACCTGAGAGTATTTCTTAATGAACCTTAAAGA 1152
 Db 909 AAATGGCATATGACATCTATTTGATGAACCTGAGAGTATTTCTTAATGAACCTTAAAGA 968
 QY 1153 TCTTAATAATCAGTGTATATGAGTGTATTCAGATTCCTCCAGTAAATTTTCAAGGGATC 1212
 Db 969 TCTTAATAATCAGTGTATATGAGTGTATTCAGATTCCTCCAGTAAATTTTCAAGGGATC 1028

```

Oy 1213 AGCCGTGTATGTATAGATGATGATGTGAGAGGGTGTCTTGTGTCATATGCCA 1272
Db 1029 AGCCGTGTATGTATAGATGATGATGTGAGAGGGTGTCTTGTGTCATATGCCA 1088
Oy 1273 CAGGATGAGCCCACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGGCC 1332
Db 1089 CAGGATGAGCCCACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGGCC 1148
Oy 1333 AGGAATCTGTCACCAAAACATTTGGTGTGCTCTCAAAAGAGCTTCCGATGA 1392
Db 1149 AGGAATCTGTCACCAAAACATTTGGTGTGCTCTCAAAAGAGCTTCCGATGA 1208
Oy 1393 TGTATTAACCTTTGAGAGAGTCAATCCAGCATGATCAATCCAGTTCCTATGA 1452
Db 1209 TGTATTAACCTTTGAGAGAGTCAATCCAGCATGATCAATCCAGTTCCTATGA 1268
Oy 1453 TCGCCCAATAGTATCAAAAAGATGTAAATTAATCAATTTACAAATTTGCTG 1512
Db 1269 TCGCCCAATAGTATCAAAAAGATGTAAATTAATCAATTTACAAATTTGCTG 1328
Oy 1513 AGTGATCAGAGATGACAGATGATGATTAATTTTCCGAGAGAGATGTTGGAG 1572
Db 1329 AGTGATCAGAGATGACAGATGATGATTAATTTTCCGAGAGAGATGTTGGAG 1388
Oy 1573 TCTTAAGTATGTTCAATTCCTAAGAGAGCTTGATGATTTAGAGAGTTCCTG 1632
Db 1389 TCTTAAGTATGTTCAATTCCTAAGAGAGCTTGATGATTTAGAGAGTTCCTG 1448
Oy 1633 AGAAATGACAGTTTTCGAGAGCCGAGCTGTAATTTTCAAGCAATGAGCTTCC 1692
Db 1449 AGAAATGACAGTTTTCGAGAGCCGAGCTGTAATTTTCAAGCAATGAGCTTCC 1508
Oy 1693 GCAACAATATATATATGTTTCAACGGCTGGGTTCGCCAGCTCCCTTACACCG 1752
Db 1509 GCAACAATATATATATGTTTCAACGGCTGGGTTCGCCAGCTCCCTTACACCG 1568
Oy 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCCCGAGAGCTTACGTGCT 1812
Db 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCCCGAGAGCTTACGTGCT 1628
Oy 1813 TGTGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGAGCGACAGAGCAAG 1872
Db 1629 TGTGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGAGCGACAGAGCAAG 1688
Oy 1873 AAGAAATGAGACCCACTGATCTATGATGTAGAGAAATGTAGCAATTTTGA 1932
Db 1689 AAGAAATGAGACCCACTGATCTATGATGTAGAGAAATGTAGCAATTTTGA 1748
Oy 1933 CAGCCCTGAGAGAGATCATCTATGATGTAGAGAAATGTAGCAATTTTGA 1992
Db 1749 CAGCCCTGAGAGAGATCATCTATGATGTAGAGAAATGTAGCAATTTTGA 1808
Oy 1993 TCCGAAAGTCGAGAGAGCGTGTGCTATTTGCAATTTCCAGAGCGCAAA 2052
Db 1809 TCCGAAAGTCGAGAGAGCGTGTGCTATTTGCAATTTCCAGAGCGCAAA 1868
Oy 2053 AGAAGAGATCAGAGTGTATGATCATATCATCAGAGCAAGTCAAGGCTTCTG 2112
Db 1869 AGAAGAGATCAGAGTGTATGATCATATCATCAGAGCAAGTCAAGGCTTCTG 1928
Oy 2113 TCTACAAAGAGAGATTCAGGCAATTAACCTGCGCATGGGGGAAACATGGG 2172
Db 1929 TCTACAAAGAGAGATTCAGGCAATTAACCTGCGCATGGGGGAAACATGGG 1988
Oy 2173 AACTCTTCTTAAGGTAACTCGAAGTCAATGACACAGAGCAATTTGAAGA 2232
Db 1989 AACTCTTCTTAAGGTAACTCGAAGTCAATGACACAGAGCAATTTGAAGA 2048
Oy 2233 TAAAGATGATGATGAGTGTCTTAAGAGCCAAAGAAATGTCCATAGACACCT 2292
Db 2049 TAAAGATGATGATGAGTGTCTTAAGAGCCAAAGAAATGTCCATAGACACCT 2108

```

```

Oy 2293 CCAGAGGTCTGTACAGAGACTTCATGACGCTCATCAACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGTACAGAGACTTCATGACGCTCATCAACCCCAATCTCAACAGAT 2168
Oy 2253 GGATGATGTTCTGTAAACAAGTTTGGAAAAGGACCCGAAAACAACTGTGG 2412
Db 2169 GGATGATGTTCTGTAAACAAGTTTGGAAAAGGACCCGAAAACAACTGTGG 2228
Oy 2413 ACATACCCAGAGGAAACAGTAAACAATGGAAGCACTTCAAGAAATATAGAAAG 2472
Db 2229 ACATACCCAGAGGAAACAGTAAACAATGGAAGCACTTCAAGAAATATAGAAAG 2288
Oy 2473 CAGAGAGCCCAAGTAATTTGAGAGGCAACCGAGAGTGTGAGCTCATTACCT 2532
Db 2289 CAGAGAGCCCAAGTAATTTGAGAGGCAACCGAGAGTGTGAGCTCATTACCT 2348
Oy 2533 AACTCAACAGAGTAAAGAACTTGCTGTGACATTAATCTGGAAGAAACAAAT 2592
Db 2349 AACTCAACAGAGTAAAGAACTTGCTGTGACATTAATCTGGAAGAAACAAAT 2408
Oy 2593 GAATTTTTTCATGACATTAATGATGATGTTTCAAAATGAGGAAATCAGCTGAG 2652
Db 2409 GAATTTTTTCATGACATTAATGATGATGTTTCAAAATGAGGAAATCAGCTGAG 2468
Oy 2653 CCAATTAATTAATTAATCAATGAGTAACTTCTTAATAGGCTTTTTT 2700
Db 2469 CCAATTAATTAATTAATCAATGAGTAACTTCTTAATAGGCTTTTTT 2516

```

RESULT 9

US-08-488-935-1

Sequence 1, Application US/08488935

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.
 APPLICANT: Kolodkin, Alex L.
 APPLICANT: Matches, David R.
 APPLICANT: Bentley, David R.
 APPLICANT: O'Connor, Timothy
 TITLE OF INVENTION: Semaphorins
 NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
 STREET: 268 Bush Street, Suite 3200
 CITY: San Francisco
 STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,935

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/489,057

FILING DATE: 09-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-2/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-08-488-935-1

Query Match 92.6%; Score 2508; DB 9; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 193 CTGACACATGGGCTGGTAACTAGATGTCTGTCTTTTCTGGGAGTATTAATTACACAC 252
DB 9 CTGACACATGGGCTGGTAACTAGATGTCTGTCTTTTCTGGGAGTATTAATTACACAC 68
QY 253 AAGAGCAAACTATCAGAAATGGAAGAAACAATGTGCCAAGGCTGAATTAATCTACAAAG 312
DB 69 AAGAGCAAACTATCAGAAATGGAAGAAACAATGTGCCAAGGCTGAATTAATCTACAAAG 128
QY 313 AATGTTGAATCCAAACATGTGATCACTTTCAATGGCTTGCCACAGCTCCAGTTATCA 372
DB 129 AATGTTGAATCCAAACATGTGATCACTTTCAATGGCTTGCCACAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTTGAAGAGAGAGTGTGTGTATGTTGAGCAAAAGATCAATAT 432
DB 189 TACCTTCTTTTGAAGAGAGAGTGTGTGTATGTTGAGCAAAAGATCAATAT 248
QY 433 TTCAATCGACCTGTGTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTAC 492
DB 249 TTCAATCGACCTGTGTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTAC 308
QY 493 CAGAGAGATGAATGCAAGTGGGCTGGAAGAAAGCATCTGAAAGAAATGTGCTAATTCAT 552
DB 309 CAGAGAGATGAATGCAAGTGGGCTGGAAGAAAGCATCTGAAAGAAATGTGCTAATTCAT 368
QY 553 CAAAGGATCTTAAGGCAATATATCAAGCTGTGACGCTGTGAGCAAGGAGCTTTTCA 612
DB 369 CAAAGGATCTTAAGGCAATATATCAAGCTGTGACGCTGTGAGCAAGGAGCTTTTCA 428
QY 613 TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAACTGGA 672
DB 429 TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAACTGGA 488
QY 673 GAACCTCAATTTTGAAGACGGCGGTGGAGAGTCCATATGACCTTAAGCTGTGACAC 732
DB 489 GAACCTCAATTTTGAAGACGGCGGTGGAGAGTCCATATGACCTTAAGCTGTGACAC 548
QY 733 ATCCCTTTAATAGATGAGATTAATCTGTGAATCTGACAGCTGATTTTATGAGGAGAGA 792
DB 549 ATCCCTTTAATAGATGAGATTAATCTGTGAATCTGACAGCTGATTTTATGAGGAGAGA 608
QY 793 CTTTGCTATCTTCCGAACTCTTGGGACCAACCACCAATCAGAGCAGACATGATTC 852
DB 609 CTTTGCTATCTTCCGAACTCTTGGGACCAACCACCAATCAGAGCAGACATGATTC 668
QY 853 CAGGTGGCTCAATGATCAAGTTCATTAAGTCCCACTCATCTCAGAGTGAACAATCC 912
DB 669 CAGGTGGCTCAATGATCAAGTTCATTAAGTCCCACTCATCTCAGAGTGAACAATCC 728
QY 913 TGAAGATGACAAAGTATTAATTTTCTTCCGTGAAATGCAATAGATGAGAACACTCTGG 972
DB 729 TGAAGATGACAAAGTATTAATTTTCTTCCGTGAAATGCAATAGATGAGAACACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGATAGGTGATGATGCAAGATGATGATTTGAGGGGACAGAG 1032
DB 789 AAAAGCTACTCAGCTAGATAGGTGATGATGCAAGATGATGATTTGAGGGGACAGAG 848
QY 1033 TCTGTGATTAATGAGCAACAATTCCTCAAGCTGTGATTTGCTCAGTCCAGGTCC 1092
DB 849 TCTGTGATTAATGAGCAACAATTCCTCAAGCTGTGATTTGCTCAGTCCAGGTCC 908
QY 1093 AAATGCAATGACATCTATTTTGAATGATGAGATGATTTCTTAATGAATCTTTAAAGA 1152
DB 909 AAATGCAATGACATCTATTTTGAATGATGAGATGATTTCTTAATGAATCTTTAAAGA 968
```

```
QY 1153 TCTTAATAATCCAGTTGATATGAGATGTTTACAGACTTCCAGTAAATTTTCAAGGATC 1212
DB 969 TCTTAATAATCCAGTTGATATGAGATGTTTACAGACTTCCAGTAAATTTTCAAGGATC 1028
QY 1213 ACCCGTGTATATGATAGATAGATGATGAGAGGTTTCTTGGTCCATATGCCCA 1272
DB 1029 ACCCGTGTATATGATAGATAGATGATGAGAGGTTTCTTGGTCCATATGCCCA 1088
QY 1273 CAGGATGAGACCAATCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1332
DB 1089 CAGGATGAGACCAATCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1148
QY 1333 AGGAATCTGTCCAGCAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1392
DB 1149 AGGAATCTGTCCAGCAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1208
QY 1393 TGTATTAACCTTTGCAAGAGATCATCCAGATGATGATGATGATGATGATGATGATGAT 1452
DB 1209 TGTATTAACCTTTGCAAGAGATCATCCAGATGATGATGATGATGATGATGATGATGAT 1268
QY 1453 TGGCCCAATGATGATCAAAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1512
DB 1269 TGGCCCAATGATGATCAAAAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1328
QY 1513 AGTGAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
DB 1329 AGTGAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
QY 1573 TCTTAAGTATGTTCAATTTCTTAAGAGACCTTGTATGATGATGATGATGATGATGAT 1632
DB 1389 TCTTAAGTATGTTCAATTTCTTAAGAGACCTTGTATGATGATGATGATGATGATGAT 1448
QY 1633 AGAATGACAGTTTGGGAAACGACCTGATGATGATGATGATGATGATGATGATGATGAT 1692
DB 1449 AGAATGACAGTTTGGGAAACGACCTGATGATGATGATGATGATGATGATGATGATGAT 1508
QY 1693 GCACCAATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
DB 1509 GCACCAATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
QY 1753 TATTTACGGAAGAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
DB 1569 TATTTACGGAAGAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
QY 1813 TGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
DB 1629 TGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
QY 1873 AAGAAATGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
DB 1689 AAGAAATGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
QY 1933 CAGCCCTGAAGAGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
DB 1749 CAGCCCTGAAGAGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
QY 1993 TCCGAAGTCCGAGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2052
DB 1809 TCCGAAGTCCGAGAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1868
QY 2053 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
DB 1869 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
QY 2113 TCTACACAGAGAGATTCAGCAATTAATCTGTGATGATGATGATGATGATGATGATGAT 2172
DB 1929 TCTACACAGAGAGATTCAGCAATTAATCTGTGATGATGATGATGATGATGATGATGAT 1988
QY 2173 AACTCTTTTAAAGTAACTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2232
DB 1989 AACTCTTTTAAAGTAACTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2048
```


QY	2233	TTAAGATATATATGAGATGCGCTCTAAGACCAAAAGATGTCCATATAGATACACTAG	2295
Db	2049	TTAAGATATATATGAGATGCGCTCTAAGACCAAAAGATGTCCATATAGATACACTAG	2106
QY	2293	CCAGAGGCTCTGTATCAGAGACTTCATGACGCTATCAACCCCAATCTTCAACAGAT	2355
Db	2109	CCAGAGGCTCTGTATCAGAGACTTCATGACGCTATCAACCCCAATCTTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGACAAAGTTGGAAAAGGACCGAAAACAAGTCGGCAAAAGGCCAGG	2412
Db	2169	GGATGAGTCTGTGACAAAGTTGGAAAAGGACCGAAAACAAGTCGGCAAAAGGCCAGG	2222
QY	2413	ACATATCCCAAGGAAACAGTAAACAATGAGAGCACTTACAGAAAAATAAGAAAGTATGAAA	2472
Db	2229	ACATATCCCAAGGAAACAGTAAACAATGAGAGCACTTACAGAAAAATAAGAAAGTATGAAA	2288
QY	2473	CAGGAGGACCCACCAATTTGAGAGGGGACCCACGAGGTCTTGAGCTGCATTACTCTTGA	2533
Db	2289	CAGGAGGACCCACCAATTTGAGAGGGGACCCACGAGGTCTTGAGGTGCAATTACTCTTGA	2348
QY	2533	AACCTCAAAACAAGTATGAAACTTGCCTTAACAATACTGGAAGAAAAACAATGCAATATACAT	2595
Db	2349	AACCTCAAAACAAGTATGAAACTTGCCTTAACAATACTGGAAGAAAAACAATGCAATATACAT	2406
QY	2593	GAACTTTTTTCATGGCATTTATGTGATGTTTCAATGTGTGGAAATTCAGCTGAGTTCCA	2655
Db	2409	GAACTTTTTTCATGGCATTTATGTGATGTTTCAATGTGTGGAAATTCAGCTGAGTTCCA	2466
QY	2653	CCAAATTAAATTAATTCATGAGAGTAACCTTCCATATAGGCTTTT	2700
Db	2469	CCAAATTAAATTAATTCATGAGAGTAACCTTCCATATAGGCTTTT	2516

RESULT 10
 US-08-489-057-1
 Sequence 1, Application US/08489057
 GENERAL INFORMATION:
 APPLICANT: Goodman, Corey S.
 APPLICANT: Kolodkin, Alex L.
 APPLICANT: Mathes, David D.
 APPLICANT: Bentley, David R.
 APPLICANT: O'Connor, Timothy
 TITLE OF INVENTION: Semaphorins
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
 STREET: 268 Bush Street, Suite 3200
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/489,057
 FILING DATE: 09-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: B94-002-2/RAO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 343-4341
 TELEFAX: (415) 343-4342
 TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2601 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: CDNA
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 16..2331
US-08-489-057-1

```

Query Match	92.6%;	Score 2508;	DB 9;	length 2601;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2508;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	193	CTGAGAGATGGCGTGGTAACTAGAGATGTCGTCTTTTCGGGGAGATTAACTTACAGC	252
Db	9	CTGAGAGATGGCGTGGTAACTAGAGATGTCGTCTTTTCGGGGAGATTAACTTACAGC	68
Qy	253	AAGAGCAACTATAGAAATGGAAAGACAAATGTCCAAGGCTGAATTAATCTCTACAAAGA	312
Db	69	AAGAGCAACTATAGAAATGGAAAGAAACAAATGTCCAAGGCTGAATTAATCTCTACAAAGA	128
Qy	313	AATGTGGAATCCAAACATGATCACTTTCAATGGGCTTGGCCAAACAGCTCAGTTATCA	372
Db	129	AATGTGGAATCCAAACATGATCACTTTCAATGGGCTTGGCCAAACAGCTCAGTTATCA	188
Qy	373	TACCTTCCTTTGGATGAGAACCGAGTAGGCTGATGTTGGAGCAAGATCAATAT	432
Db	189	TACCTTCCTTTGGATGAGAACCGAGTAGGCTGATGTTGGAGCAAGATCAATAT	248
Qy	433	TTCAATTGACCTGGTTAAATCAAGATTTTCAAAAAGTTGTGTCGCAATATCTTAAAC	492
Db	249	TTCAATTGACCTGGTTAAATCAAGATTTTCAAAAAGTTGTGTCGCAATATCTTAAAC	308
Qy	493	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGATCCTGAAAGATGTGTAATTTGAT	552
Db	309	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGATCCTGAAAGATGTGTAATTTGAT	368
Qy	553	CAAGTACTTAAGGCATATATACAGACTCACTTGTACGCTGTGGAAAGGGGGCTTTTCA	612
Db	369	CAAGTACTTAAGGCATATATACAGACTCACTTGTACGCTGTGGAAAGGGGGCTTTTCA	428
Qy	613	TTCGAATTTGACCTACACATTGAAATTTGACATCATCTGAGGACAAATATTTTAACTGGA	672
Db	429	TTCGAATTTGACCTACACATTGAAATTTGACATCATCTGAGGACAAATATTTTAACTGGA	488
Qy	673	GAACTCACATTTTGAAGAACGGCCGTGGGAAAGATCCATATGACCTTAAGCTGTGACAGC	732
Db	489	GAACTCACATTTTGAAGAACGGCCGTGGGAAAGATCCATATGACCTTAAGCTGTGACAGC	548
Qy	733	ATCCCTTTTAATAGATGGAATTTATACTGGAACCTGACAGCTGATTTTATGGGCGAGA	792
Db	549	ATCCCTTTTAATAGATGGAATTTATACTGGAACCTGACAGCTGATTTTATGGGCGAGA	608
Qy	793	CTTTGCATTTTCCGAATCTTTGGGGACACACACCCCAATCAGACACAGCAGCATATTC	852
Db	609	CTTTGCATTTTCCGAATCTTTGGGGACACACACCCCAATCAGACACAGCAGCATATTC	668
Qy	853	CAGGTGCTCAATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGATGACAAATCC	912
Db	669	CAGGTGCTCAATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGATGACAAATCC	728
Qy	913	TGAAGATGACAAAGTATACCTTTTCTTCCGTGAATAATGCAATAGATGAGAACCTCTGG	972
Db	729	TGAAGATGACAAAGTATACCTTTTCTTCCGTGAATAATGCAATAGATGAGAACCTCTGG	788
Qy	973	AAAAAGTACACGAGCTGGAATAAGGTGCAATATGCAAGAAATGACCTTGGAGGGCACAGAG	1032
Db	789	AAAAAGTACACGAGCTGGAATAAGGTGCAATATGCAAGAAATGACCTTGGAGGGCACAGAG	848
Qy	1033	TCTGTGGAATTAATGCAACAATTCCTCAAGAGCTGCTGTGATTTGCTCAGTGCAGAGTCC	1092
Db	849	TCTGTGGAATTAATGCAACAATTCCTCAAGAGCTGCTGTGATTTGCTCAGTGCAGAGTCC	908
Qy	1093	AAATGGCATTGACACTATTTTGTATGAATGACAGAGATGATTTCTTAATGAACCTTTAAGA	1152

Db 909 AATGACATTCACATCTATTGATGAACTGACAGATGATATTCCTAATGAACCTTTAAGA 968
Qy 1153 TCTTAATAATCCAGTTGTATATGAGAGTTTACAGCTCCAGTACATTTTCAAGGAGTC 1212
Db 969 TCTTAATAATCCAGTTGTATATGAGAGTTTACAGCTCCAGTACATTTTCAAGGAGTC 1028
Qy 1213 AGCCGTGTATATGATAGATGATGATGAGAGGTTTCTTGTTCATATGCCCCA 1272
Db 1029 AGCCGTGTATATGATAGATGATGATGAGAGGTTTCTTGTTCATATGCCCCA 1088
Qy 1273 CAGGATGAGACCAATCATATGAGTCTTTCAGAGAGAGTCCCTATCCAGGACC 1332
Db 1089 CAGGATGAGACCAATCATATGAGTCTTTCAGAGAGAGTCCCTATCCAGGACC 1148
Qy 1333 AGGAACTTGTCCAGCAAAACATTTGTGTTTGACTCTCAAAAGAACCTTCTGATGA 1392
Db 1149 AGGAACTTGTCCAGCAAAACATTTGTGTTTGACTCTCAAAAGAACCTTCTGATGA 1208
Qy 1393 TGTATATACCTTTGCAAGAGTGTATGATGATGATGATGATGATGATGATGATGAT 1452
Db 1209 TGTATATACCTTTGCAAGAGTGTATGATGATGATGATGATGATGATGATGATGAT 1268
Qy 1453 TGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
Db 1269 TGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
Qy 1513 AGTGAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db 1329 AGTGAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
Qy 1573 TCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
Db 1389 TCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
Qy 1633 AGAATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
Db 1449 AGAATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
Qy 1693 GCAACACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
Db 1509 GCAACACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
Qy 1753 TATTTAAGGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
Db 1569 TATTTAAGGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
Qy 1813 TGTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
Db 1629 TGTGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
Qy 1873 AAGAAATGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
Db 1689 AAGAAATGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
Qy 1933 CAGCCCTGAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
Db 1749 CAGCCCTGAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
Qy 1993 TCCGAAGTGCAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2052
Db 1809 TCCGAAGTGCAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
Qy 2053 AGAAGAGATCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2112
Db 1869 AGAAGAGATCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
Qy 2113 TCTAACAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2172
Db 1929 TCTAACAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1988
Qy 2173 AACTCTTTTAAAGTAACTTGAAGTCAATGACAGAGCAATTTGGAAGAACTTTTCA 2232
Db 1989 AACTCTTTTAAAGTAACTTGAAGTCAATGACAGAGCAATTTGGAAGAACTTTTCA 2048

Qy 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
Db 2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108
Qy 2293 CAGAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
Db 2109 CAGAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2168
Qy 2353 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2412
Db 2169 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2228
Qy 2413 ACATACCCAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2472
Db 2229 ACATACCCAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2288
Qy 2473 CAGAGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2532
Db 2289 CAGAGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2348
Qy 2533 AACCTCAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2592
Db 2349 AACCTCAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2408
Qy 2593 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
Db 2409 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
Qy 2653 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2710
Db 2469 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2516

RESULT 11

US-10-067-632-53

Sequence 53, Application US/10067632

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

Kolodkin, Alex L.

Metthes, David R.

Bentley, David R.

O'Connor, Timothy

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESSES:

ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 3200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/067,632

FILING DATE: 04-Feb-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/060,610

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/835,268

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Oseman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342

TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-067-632-53

Query Match 92.6%; Score 2508; DB 45; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 CTGAGCATGGCTGGTTAACTAGGATTTCTCTTTCTGGGAGATTAATTACAGC 252
9 CTGAGCATGGCTGGTTAACTAGGATTTCTCTTTCTGGGAGATTAATTACAGC 68

253 AAGACCAACTATCGAATGGGAAGAACAATGGCCAGGCTGAATTAATCTTACAAAG 312
69 AAGACCAACTATCGAATGGGAAGAACAATGGCCAGGCTGAATTAATCTTACAAAG 128

313 AATGTGGAACTCAACATGTGATCACTTCAATGGCTGGCCACAGCTCCAGTTATCA 372
129 AATGTGGAACTCAACATGTGATCACTTCAATGGCTGGCCACAGCTCCAGTTATCA 188

373 TACCTTCCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGATCACATATT 432
189 TACCTTCCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGATCACATATT 248

433 TTCAATTCGACCTGGTTAATATCAAGATTTTCAAAAGATTTGTGTGCCAGTATCTTAAC 492
249 TTCAATTCGACCTGGTTAATATCAAGATTTTCAAAAGATTTGTGTGCCAGTATCTTAAC 308

493 CAGAAGATGGAATCAAGTGGCTGGAAGAAAGACATCTGAAAGATGTCTAATTTTCA 552
309 CAGAAGATGGAATCAAGTGGCTGGAAGAAAGACATCTGAAAGATGTCTAATTTTCA 368

553 CAAGGTACTTAAAGCATTAATCAAGCTCACTTGAAGCTGTGGACGGGGCTTTTCA 612
369 CAAGGTACTTAAAGCATTAATCAAGCTCACTTGAAGCTGTGGACGGGGCTTTTCA 428

613 TTCAATTTGCACTTACATTTGAATTTGACATCATCTTGAGACATATTTTAACTGGA 672
429 TTCAATTTGCACTTACATTTGAATTTGACATCATCTTGAGACATATTTTAACTGGA 488

673 GAACCTCACTTTGAAACGGCGGTGGGAAGAGTCATATGACCTTAAGCTGTGACAGC 732
489 GAACCTCACTTTGAAACGGCGGTGGGAAGAGTCATATGACCTTAAGCTGTGACAGC 548

733 ATCCCTTTAATAGATGAGATTAATTAATCTGGAAGCTGAGCTGATTTTATGGGCGAGA 792
549 ATCCCTTTAATAGATGAGATTAATTAATCTGGAAGCTGAGCTGATTTTATGGGCGAGA 608

793 CTTTGCATCTTGGGAACCTTGGGACACCAACCAATCAGAGACAGAGCATGATTC 852
609 CTTTGCATCTTGGGAACCTTGGGACACCAACCAATCAGAGACAGAGCATGATTC 668

853 CAGGTGGCTCAATGATCCAAAGTTGATAGTGGCCACCTCATCTGAGAGTGAACAATCC 912
669 CAGGTGGCTCAATGATCCAAAGTTGATAGTGGCCACCTCATCTGAGAGTGAACAATCC 728

913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGACACTGCG 972
729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGACACTGCG 788

973 AAAAGTACTACGCTAGATAGTGCATATGCAAGATGACTTTGAGGGGACAGAG 1032
789 AAAAGTACTACGCTAGATAGTGCATATGCAAGATGACTTTGAGGGGACAGAG 848

1033 TCTGTGATTAATAGGACAAATCTCTCAAGCTGCTGTGATTTGCTGAGCCAGGTCC 1092
849 TCTGTGATTAATAGGACAAATCTCTCAAGCTGCTGTGATTTGCTGAGCCAGGTCC 908

1093 AAATGGCACTGACATCTATTTGATGAATGCAAGATGATTTCTTAATGAATTTAAGA 1152
909 AAATGGCACTGACATCTATTTGATGAATGCAAGATGATTTCTTAATGAATTTAAGA 968

1153 TCTTAATAATCAGTGTATATGAGTGTATACACTTCCAGTAACATTTCAAGGGATC 1212
969 TCTTAATAATCAGTGTATATGAGTGTATACACTTCCAGTAACATTTCAAGGGATC 1028

1213 AGCGTGTATATGATAGCATGATGATGATGAGAGGAGTGTCTTGGTCCATATGCCCA 1272
1029 AGCGTGTATATGATAGCATGATGATGAGAGGAGTGTCTTGGTCCATATGCCCA 1088

1273 CAGGATGACCCCAATATCAATGGGTGCTTATCAAGAAAGTCCCTATCAAGGCG 1332
1089 CAGGATGACCCCAATATCAATGGGTGCTTATCAAGAAAGTCCCTATCAAGGCG 1148

1333 AGGAATCTGTCCAGCAAAACATTTGGTGTGATGACTTACAAAGACCTTCTGATGA 1392
1149 AGGAATCTGTCCAGCAAAACATTTGGTGTGATGACTTACAAAGACCTTCTGATGA 1208

1393 TGTATAACCTTTGCAAGAAATCATCCAGCATGTAACATGATTTCTTATGAACAA 1452
1209 TGTATAACCTTTGCAAGAAATCATCCAGCATGTAACATGATTTCTTATGAACAA 1268

1453 TCGCCCAATAGTATCAAAACGAGATTAATTAATCAATTAACAAATTTGTGTAGACCG 1512
1269 TCGCCCAATAGTATCAAAACGAGATTAATTAATCAATTAATCAATTTGTGTAGACCG 1328

1513 AGTGATGCAAAATGAGACAGTATGATGATTTATGGAACAAATGTTGGGACCGT 1572
1329 AGTGATGCAAAATGAGACAGTATGATGATTTATGGAACAAATGTTGGGACCGT 1388

1573 TCTTAAGTATTTCAATTTCTTAAGAGACTTGTATGATTTAAGAGTGTCTGCTGGA 1632
1389 TCTTAAGTATTTCAATTTCTTAAGAGACTTGTATGATTTAAGAGAGTGTCTGCTGGA 1448

1633 AGAATGACAGTTTTTGGGAAACGAGCTGATTTTCAAGCAATGAGCTTCCACTAAGCA 1692
1449 AGAATGACAGTTTTTGGGAAACGAGCTGATTTTCAAGCAATGAGCTTCCACTAAGCA 1508

1693 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGA 1752
1509 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGA 1568

1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGAACCTTACTGTGCTGGA 1812
1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGAACCTTACTGTGCTGGA 1628

1813 TGGTCTGCAATTTCTGCTATTTTCCACTGCAAGAGAGCAAGACGACAAGATAT 1872
1629 TGGTCTGCAATTTCTGCTATTTTCCACTGCAAGAGAGCAAGACGACAAGATAT 1688

1873 AAGAAATGAGACCCCACTGATCTGCTTCACTTACCATGATATATCACTATGCGCA 1932
1689 AAGAAATGAGACCCCACTGATCTGCTTCACTTACCATGATATATCACTATGCGCA 1748

1933 CAGCCTGGAAGAGATCATCTATGAGTGTGAGAAATAGTACATTTTGGAAATGAG 1992
1749 CAGCCTGGAAGAGATCATCTATGAGTGTGAGAAATAGTACATTTTGGAAATGAG 1808

1993 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGGCAATGAAGCGAAA 2052
1809 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGGCAATGAAGCGAAA 1868

2053 AGAAGATGATGATGATGATATCATATCATAGAGACAGATTAAGGCTTCTGCTACGTAG 2112
1869 AGAAGATGATGATGATGATATCATATCATAGAGACAGATTAAGGCTTCTGCTACGTAG 1928

QY	2113	TCCTAACAAGAAAGATTTCAGGCAATTACTCTCGGCCATCGGGGAAACATGGGTTTCATCA	2172
Db	1929	TCCTAACAAGAAAGATTTCAGGCAATTACTCTCGGCCATCGGGGAAACATGGGTTTCATCA	1988
QY	2173	AACTCTTCTTAAGGTAAACCTCGGAAGTCATTGACACAGAGCATTTGGAGAAGACTTCTTCA	2232
Db	1989	AACTCTTCTTAAGGTAAACCTCGGAAGTCATTGACACAGAGCATTTGGAGAAGACTTCTTCA	2048
QY	2233	TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCTAATAGCATGACCTAG	2292
Db	2049	TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCTAATAGCATGACCTAG	2108
QY	2293	CCAGAAGTCGTGATCAGAGACTTATGACAGTATCAACCAACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAAGTCGTGATCAGAGACTTATGACAGTATCAACCAACCCCAATCTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGTAACAAGTTTGAAAAAGGACCGAANAACATCTCGGCAAAAGGCGCAG	2412
Db	2169	GGATGAGTCTGTGTAACAAGTTTGAAAAAGGACCGAANAACATCTCGGCAAAAGGCGCAG	2228
QY	2413	ACATACCCCAAGGAAACAGTAACAAATGGAAGCATTTACAAGAAATTAAGAAAGGTAGAA	2472
Db	2229	ACATACCCCAAGGAAACAGTAACAAATGGAAGCATTTACAAGAAATTAAGAAAGGTAGAA	2288
QY	2473	CAGAGAGAACCCACGAATTTGAGAGGGCACCCAGAGGTCTTGAGCTGATTAACCTCTGA	2532
Db	2289	CAGAGAGAACCCACGAATTTGAGAGGGCACCCAGAGGTCTTGAGCTGATTAACCTCTGA	2348
QY	2533	AACCTCAACAAGTAAAGACTGCTTAGACATAATCTGAAAAAACAATGCAATTATCAT	2592
Db	2349	AACCTCAACAAGTAAAGACTGCTTAGACATAATCTGAAAAAACAATGCAATTATCAT	2408
QY	2593	GAACTTTTTCATGGCATTTATGTGATGTTTACAATGATGGGAAATTCAGCTGAGTTCCA	2652
Db	2409	GAACTTTTTCATGGCATTTATGTGATGTTTACAATGATGGGAAATTCAGCTGAGTTCCA	2468
QY	2653	CCAATTATTAATTAATCCATGACATTAATCTTCCATATAGGCTTTT	2710
Db	2469	CCAATTATTAATTAATCCATGACATTAATCTTCCATATAGGCTTTT	2516

```

RESULT 12
US-10-170-235-19890
; Sequence 19890, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CLO01380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 19890
; LENGTH: 2857
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-19890

      82.2%   Score 2226; DB 47; Length 2857;
Query Match Best Local Similarity 99.8%; Pred No. 0;
Matches 2516; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Oy      193  CTGCAGCATGCGCTGCTTAACTAGGATTTGTCCTCTTTCTGGGGAGTATTACTTACAGC 252
Db      1    CTGCAGCATGGGCTGCTTAACTAGGATTTGTCCTCTTTCTGGGGAGTATTACTTACAGC 60
Oy      253  AAGAGCAACTATCAGAAATGGGAGAACATATGCCCAAGGCTGAAATATCTCTACAAGA 312
Db      61  AAGAGCAACTATCAGAAATGGGAGAACATATGCCCAAGGCTGAAATATCTCTACAAGA 120
Oy      313  AATGTGGAATCCAACAATGTGATCATTTCATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db      121  AATGTGGAATCCAACAATGTGATCATTTCATGGCTTGGCCAAACAGCTCCAGTTATCA 180

```

QY	373	TACCTTCCTTTGGATGAGGAACGAGTAGCCTGTATGTTGGACCAAGATCAATATT	432
Db	181	TACCTTCCTTTGGATGAGGAACGAGTAGCCTGTATGTTGGACCAAGATCAATATT	240
QY	433	TTCAATCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	492
QY	493	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGACATCTTGAAAGAAATGTGTATTTTCAT	552
Db	301	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGACATCTTGAAAGAAATGTGTATTTTCAT	360
QY	553	CAAGGTACTTAAGGCATTAATACGACTCACTTGTACGCCCTGTGAAACGGGGGCTTTTCA	612
Db	351	CAAGGTACTTAAGGCATTAATACGACTCACTTGTACGCCCTGTGAAACGGGGGCTTTTCA	420
QY	613	TCCAATTTGCACTTACATTTGAATTTGGACATCATCTTGAGGACAAATATTTTAAAGCTGGA	672
Db	421	TCCAATTTGCACTTACATTTGAATTTGGACATCATCTTGAGGACAAATATTTTAAAGCTGGA	480
QY	673	GAACTCAATTTTAAAAACGGCCGTGGGAAGATGCATATGACCCCTAAGCTGACAGC	732
Db	481	GAACTCAATTTTAAAAACGGCCGTGGGAAGATGCATATGACCCCTAAGCTGACAGC	540
QY	733	ATCCCTTTAATAGATGGAGAAATTAATCTGTGAATGTGCACTGATTTTATGGGGCGAGA	792
Db	541	ATCCCTTTAATAGATGGAGAAATTAATCTGTGAATGTGCACTGATTTTATGGGGCGAGA	600
QY	793	CTTTGCTATCTTTCCGAACCTTTGGGCACACACCACCAATCAGACACAGCAGCATGATTC	852
Db	601	CTTTGCTATCTTTCCGAACCTTTGGGCACACACCACCAATCAGACACAGCAGCATGATTC	660
QY	853	CAGGTGGCTCATGATCCAAAGTTCAATTAAGGCCACCTCATCTCAGAGATGTCAATATCC	912
Db	661	CAGGTGGCTCATGATCCAAAGTTCAATTAAGGCCACCTCATCTCAGAGATGTCAATATCC	720
QY	913	TGAAGATGACAAAGTATACCTTTTTCTTCGTGAAAATGCAATATAGTGAAGAACACTCTGG	972
Db	721	TGAAGATGACAAAGTATACCTTTTTCTTCGTGAAAATGCAATATAGTGAAGAACACTCTGG	780
QY	973	AAAAGCTATCAGCGTGAATAGGTACAGATATGCAAGATATGACTTTGGAGGGCAGAGAAG	1032
Db	781	AAAAGCTATCAGCGTGAATAGGTACAGATATGCAAGATATGACTTTGGAGGGCAGAGAAG	840
QY	1033	TCTGTGTAATAATGACCAACATTTCTTCAAAAGCTGTTGATTTTCTCAGTGCAGAGTCC	1092
Db	841	TCTGTGTAATAATGACCAACATTTCTTCAAAAGCTGTTGATTTTCTCAGTGCAGAGTCC	900
QY	1093	AAATGGAATGACACTCACTTTTGTATGAACTGCAAGATGTATCTCTAATGAACTTTAAGA	1152
Db	901	AAATGGAATGACACTCACTTTTGTATGAACTGCAAGATGTATCTCTAATGAACTTTAAGA	960
QY	1153	TCCTTAAATATCAATGTGTATATGAGAGTGTATACACTTCCAGTAAACATTTTCAAGGATC	1212
Db	961	TCCTTAAATATCAATGTGTATATGAGAGTGTATACACTTCCAGTAAACATTTTCAAGGATC	1020
QY	1213	AGCGGTGTATGTATATGACATGAGTATGTAGAAAGGGTGTCTTGTGCTCATATGCCCA	1272
Db	1021	AGCGGTGTATGTATATGACATGAGTATGTAGAAAGGGTGTCTTGTGCTCATATGCCCA	1080
QY	1273	CAGGGATGACCCAACTATCAATGGGAGCCTTATCAAGAAAGAGTCCCTATCCACGGCC	1332
Db	1081	CAGGGATGACCCAACTATCAATGGGAGCCTTATCAAGAAAGAGTCCCTATCCACGGCC	1140
QY	1333	AGGAATTTGTCCACAGCAAAACATTTGTGTGTTTGACTTACAAAGACCTTCTGTATGA	1392
Db	1141	AGGAATTTGTCCACAGCAAAACATTTGTGTGTTTGACTTACAAAGACCTTCTGTATGA	1200
QY	1393	TGTTATATACCTTTGCAAGAAATCATCCAGCCATGTATCAATCCAAATGTTTCTATGAACAA	1452
Db	1201	TGTTATATACCTTTGCAAGAAATCATCCAGCCATGTATCAATCCAAATGTTTCTATGAACAA	1260

QY 701 AAGAGTCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATAGATGAGATTTATAC 760
 Db 643 AAGAGTCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATAGATGAGATTTATAC 702
 QY 761 TCTGGAACCTGACCTGATTTTATGAGGCGAGATCTTTGCTATCTTCCGAACCTTGGGCGAC 820
 Db 703 TCTGGAACCTGACCTGATTTTATGAGGCGAGATCTTTGCTATCTTCCGAACCTTGGGCGAC 762
 QY 821 CACCAACCCATGACGACAGAGCATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 880
 Db 763 CACCAACCCATGACGACAGAGCATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 822
 QY 881 AGTGCCCATCTCATCTCAAGAGTGAACAATCTGAAAGATGACAAAGTATATCTTTTCTTC 940
 Db 823 AGTGCCCATCTCATCTCAAGAGTGAACAATCTGAAAGATGACAAAGTATATCTTTTCTTC 882
 QY 941 CCGTGAATAATGCAATATGAGAGAACCTCTGAAAAAGTACTACGCTAGATAGGTGAG 1000
 Db 883 CCGTGAATAATGCAATATGAGAGAACCTCTGAAAAAGTACTACGCTAGATAGGTGAG 942
 QY 1001 ATATGCAAGATGACCTTTGAGGGGACAGAGTCTGGTGAATTAATGACCAATCTCTC 1060
 Db 943 ATATGCAAGATGACCTTTGAGGGGACAGAGTCTGGTGAATTAATGACCAATCTCTC 1002
 QY 1061 AAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGSCATTTGACACTCATTTTGATGA 1120
 Db 1003 AAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGSCATTTGACACTCATTTTGATGA 1062
 QY 1121 CTGCGAGATGATTTCTTAATGAACTTTAAGATCCCTAATAATCCAGTTGTATATGAGAG 1180
 Db 1063 CTGCGAGATGATTTCTTAATGAACTTTAAGATCCCTAATAATCCAGTTGTATATGAGAG 1122
 QY 1181 TTACGACTTCAGTAACATTTTCAAGGATCAGCCGTGTATATGATAGCATGATGAT 1240
 Db 1123 TTACGACTTCAGTAACATTTTCAAGGATCAGCCGTGTATATGATAGCATGATGAT 1182
 QY 1241 GTGAGAGGGTGTCTTCTGCTCATATGCCCACAGGATGACCCAACTATCAATGGGTG 1300
 Db 1183 GTGAGAGGGTGTCTTCTGCTCATATGCCCACAGGATGACCCAACTATCAATGGGTG 1242
 QY 1301 CTTTATCAAGAGAGATCCCCCTATTCACAGGCGAGAACTTTGCCGCAAAACATTTGGT 1360
 Db 1243 CTTTATCAAGAGAGATCCCCCTATTCACAGGCGAGAACTTTGCCGCAAAACATTTGGT 1302
 QY 1361 GGTTTGACTCTACAAAGAGACCTTCTGATGATGATTAACCTTTSCAAGAGTCAATCA 1420
 Db 1303 GGTTTGACTCTACAAAGAGACCTTCTGATGATGATTAACCTTTSCAAGAGTCAATCA 1362
 QY 1421 GCCATGATCAATCCAGTGTCTTCTCATGAAACAATGCCCAATAGTATCAAAAAGATGTA 1480
 Db 1363 GCCATGATCAATCCAGTGTCTTCTCATGAAACAATGCCCAATAGTATCAAAAAGATGTA 1422
 QY 1481 AATTATCAATTTTACCAAAATGTCGAGCCGAGTGGATGACAGAGATGAGATGAT 1540
 Db 1423 AATTATCAATTTTACCAAAATGTCGAGCCGAGTGGATGACAGAGATGAGATGAT 1482
 QY 1541 GTTATGTTATGCAAGACAGATGTTGGACCGTTCTTAAAGTATGATTTCAATCTTAAGAG 1600
 Db 1483 GTTATGTTATGCAAGACAGATGTTGGACCGTTCTTAAAGTATGATTTCAATCTTAAGAG 1542
 QY 1601 ACTTGATA 1608
 Db 1543 ACTTGATA 1550

RESULT 14

US-09-652-816-7308
 ; Sequence 7308, Application US/09652816
 ; GENERAL INFORMATION:
 ; APPLICANT: Gutierrez-Ramos, Jose-Carlos
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.1177-001

; CURRENT APPLICATION NUMBER: US/09/652,816
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 60/152,111
 ; PRIORITY FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 9647
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7308
 ; LENGTH: 2256
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(2256)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-652-816-7308

Query Match 51.2%; Score 1388; DB 28; Length 2256;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 GTCTGCTTTTCTGAGGAGTATTAATCTTACAGCAAGCAAACTATCAGAAATGGAAGAAC 280
 Db 163 GTCTGCTTTTCTGAGGAGTATTAATCTTACAGCAAGCAAACTATCAGAAATGGAAGAAC 222
 QY 281 AATGTCAGAGGCTGAATTTATCTTACAAAGAAATGTTGAATCCAAATGTGATCACT 340
 Db 223 AATGTCAGAGGCTGAATTTATCTTACAAAGAAATGTTGAATCCAAATGTGATCACT 282
 QY 341 TTCAATGCTTGGCCCAACAGCTCCAGTTATCATACCTCTTTGGATGAGGAAGGAGT 400
 Db 283 TTCAATGCTTGGCCCAACAGCTCCAGTTATCATACCTCTTTGGATGAGGAAGGAGT 342
 QY 401 AGGCTGTATGTTGAGCAAGAGATCAATATTTTCAATGCACTGAGCTGTTAATATCAAGAT 460
 Db 343 AGGCTGTATGTTGAGCAAGAGATCAATATTTTCAATGCACTGAGCTGTTAATATCAAGAT 402
 QY 461 TTTCAAAAGATTTGTGCGCAGATATCTTACACAGAGAGATGAATGCAAGTGGGCTGGA 520
 Db 403 TTTCAAAAGATTTGTGCGCAGATATCTTACACAGAGAGATGAATGCAAGTGGGCTGGA 462
 QY 521 AAAGCATCTCGAAAGATATGCTATTTTCAATCAAGATGATCTTAAGGATATATCAAGT 580
 Db 463 AAAGCATCTCGAAAGATATGCTATTTTCAATCAAGATGATCTTAAGGATATATCAAGT 522
 QY 581 CACTGTAGCGCTGAGAAAGGAGGCTTTTATCAATCAATTTTCAATCAATTTTCAATTTTGA 640
 Db 523 CACTGTAGCGCTGAGAAAGGAGGCTTTTATCAATCAATTTTCAATCAATTTTCAATTTTGA 582
 QY 641 CATCATCTGAGAGCAATATTTTAAAGTGAAGAACTCAATTTTGAAGAGGCGCTGAG 700
 Db 583 CATCATCTGAGAGCAATATTTTAAAGTGAAGAACTCAATTTTGAAGAGGCGCTGAG 642
 QY 701 AAGATCCATATGACCCCTAAGCTGCTGACACATCCCTTTAATAGAGGAATATATAC 760
 Db 643 AAGATCCATATGACCCCTAAGCTGCTGACACATCCCTTTAATAGAGGAATATATAC 702
 QY 761 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTGCTATCTTCCGAACCTTGGGCGAC 820
 Db 703 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTGCTATCTTCCGAACCTTGGGCGAC 762
 QY 821 CACCAACCCATCAGAGCAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 880
 Db 763 CACCAACCCATCAGAGCAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 822
 QY 881 AGTGCCCATCTCATCTCAAGAGTGAACAATCTGAAAGATGACAAAGTATATCTTTTCTTC 940
 Db 823 AGTGCCCATCTCATCTCAAGAGTGAACAATCTGAAAGATGACAAAGTATATCTTTTCTTC 882
 QY 941 CCGTGAATAATGCAATATGAGAGAACCTCTGAAAAAGTACTACGCTAGATAGGTGAG 1000
 Db 883 CCGTGAATAATGCAATATGAGAGAACCTCTGAAAAAGTACTACGCTAGATAGGTGAG 942
 QY 1001 ATATGCAAGATGACCTTTGAGGGGACAGAGTCTGTGATTAATGAGCAAACTCTCTC 1060

Db 943 AATATGCAAGAAATGACTTTGGAGGACACAAAGCTGGTGAATAATGACCAATTTCC 1002
Qy 1061 AAAGCTCGCTGATTTGCTCACTGCGCAGGTCCTCAAAATGATGACACTATTTGATGA 1120
Db 1003 AAAGCTCGCTGATTTGCTCACTGCGCAGGTCCTCAAAATGATGACACTATTTGATGA 1062
Qy 1121 CTGCAAGATGATATTCCTTAATGAATTTAAAGATCTTAAATCAAGTTGATATGAGATG 1180
Db 1063 CTGCAAGATGATATTCCTTAATGAATTTAAAGATCTTAAATCAAGTTGATATGAGATG 1122
Qy 1181 TTTAGACTTCCAGTAACATTTTCAAGGATCAAGCCGTGTATGATATGATGATGAT 1240
Db 1123 TTTAGACTTCCAGTAACATTTTCAAGGATCAAGCCGTGTATGATATGATGATGAT 1182
Qy 1241 GTGAGAAGGGTGTCTTGGTCCATATGCGCAGGATGACCCAACTATCAATGGGGT 1300
Db 1183 GTGAGAAGGGTGTCTTGGTCCATATGCGCAGGATGACCCAACTATCAATGGGGT 1242
Qy 1301 CCTTATCAAGAAAGTCCCTATATCAAGGACAGGACCTTGTCCAGCAAAATTTGGT 1360
Db 1243 CCTTATCAAGAAAGTCCCTATATCAAGGACAGGACCTTGTCCAGCAAAATTTGGT 1302
Qy 1361 GGTTTGACTCTACAAAGACCTTCTGATGATGTTATTAACCTTTGCAAGAACTATCA 1420
Db 1303 GGTTTGACTCTACAAAGACCTTCTGATGATGTTATTAACCTTTGCAAGAACTATCA 1362
Qy 1421 GCCATGTCATATCAAGTGTCTTATGAAATATGCGCAATGATGATCAAAACGATGTA 1480
Db 1363 GCCATGTCATATCAAGTGTCTTATGAAATATGCGCAATGATGATCAAAACGATGTA 1422
Qy 1481 AATTATCAATTTACAAATATGTCGTAGACGAGTGTAGTACAGAAATGACATATGAT 1540
Db 1423 AATTATCAATTTACAAATATGTCGTAGACGAGTGTAGTACAGAAATGACATATGAT 1482
Qy 1541 GTTATGTTTATGGAACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTTAAGAG 1600
Db 1483 GTTATGTTTATGGAACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTTAAGAG 1542
Qy 1601 ACTTGATA 1608
Db 1543 ACTTGATA 1550

RESULT 15
US-09-652-918-7516
; Sequence 7516, Application US/09652918
; GENERAL INFORMATION:
; APPLICANT: Galvin, Katherine
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600, 1187-001
; CURRENT APPLICATION NUMBER: US/09/652, 918
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151, 130
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 8985
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7516
; LENGTH: 2256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2256)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-918-7516

Query Match 51.2%; Score 1388; DB 28; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 221 GTCTGCTTTTCTGGGGAGTATTACTTACAGCAAGACAAATATACAGATGGAGAAC 280
Db 153 GTCTGCTTTTCTGGGGAGTATTACTTACAGCAAGACAAATATACAGATGGAGAAC 222
Qy 281 AATGCGCAAGCGTAATTTATCCACAAAGAAATGTTGAAATCCAAATGATGATCT 340
Db 223 AATGCGCAAGCGTAATTTATCCACAAAGAAATGTTGAAATCCAAATGATGATCT 282
Qy 341 TTCAATGCTTGGCCAAAGCTCCAGTATATACCTTCTTTGATGAGAAACGAGT 400
Db 283 TTCAATGCTTGGCCAAAGCTCCAGTATATACCTTCTTTGATGAGAAACGAGT 342
Qy 401 AGGCTGATGTTGAGCAAGAGATCAATATTTTCAATGCAAGCTGTATATATCAAGAT 460
Db 343 AGGCTGATGTTGAGCAAGAGATCAATATTTTCAATGCAAGCTGTATATATCAAGAT 402
Qy 461 TTTCAAAAGATTTGTGGCCAGTATCTTACACAGAAAGATGATGCAAGTGGGCTGGA 520
Db 403 TTTCAAAAGATTTGTGGCCAGTATCTTACACAGAAAGATGATGCAAGTGGGCTGGA 462
Qy 521 AAAGCATCTGAAAGAAATGTGCTAATTTTCAATCAAGGTAAGGATATATACAGCT 580
Db 463 AAAGCATCTGAAAGAAATGTGCTAATTTTCAATCAAGGTAAGGATATATACAGCT 522
Qy 581 CACTTGTACGCTGTGGAACGGGGCTTTTATCAATCAATTTGCAAGCTTATGAAATTTGA 640
Db 523 CACTTGTACGCTGTGGAACGGGGCTTTTATCAATCAATTTGCAAGCTTATGAAATTTGA 582
Qy 641 CATATCTGTAGAGCAATATTTTAAAGTGTGGAATCTACATTTTGAAGAGGCGCTGGG 700
Db 583 CATATCTGTAGAGCAATATTTTAAAGTGTGGAATCTACATTTTGAAGAGGCGCTGGG 642
Qy 701 AAGAGTCATATGACCCCTAAGCTGTGACAGATCCCTTTAATATGATGAGAAATTAAC 760
Db 643 AAGAGTCATATGACCCCTAAGCTGTGACAGATCCCTTTAATATGATGAGAAATTAAC 702
Qy 761 TCTGGAATGAGCTGATTTTATGAGGGGAGACCTTGTATCTTCCGAACTCTTGGGAC 820
Db 703 TCTGGAATGAGCTGATTTTATGAGGGGAGACCTTGTATCTTCCGAACTCTTGGGAC 762
Qy 821 CACCAACCAATCAGAGACAGACAGATATTCAGAGTGGCTCAATGATCCAAAGTTCAAT 880
Db 763 CACCAACCAATCAGAGACAGACAGATATTCAGAGTGGCTCAATGATCCAAAGTTCAAT 822
Qy 881 AGTGCCCACTCATCTCAGAGATGACATCTGAAAGTGAACAAAGTATCTTTTCTTC 940
Db 823 AGTGCCCACTCATCTCAGAGATGACATCTGAAAGTGAACAAAGTATCTTTTCTTC 882
Qy 941 CGTGAAATATGCAATATGATGAGAAACATCTGAAAGTGAACAAAGTATCTTTTCTTC 1000
Db 883 CGTGAAATATGCAATATGATGAGAAACATCTGAAAGTGAACAAAGTATCTTTTCTTC 942
Qy 1001 ATATGCAAGATGACTTTGAGGGGACAGAACTGTGATGATTAATGACAACATTCCTC 1060
Db 943 ATATGCAAGATGACTTTGAGGGGACAGAACTGTGATGATTAATGACAACATTCCTC 1002
Qy 1061 AAAGCTGCTGATTTGCTCAGTGCAGGTCCTCAATGATGATGACATCTATTTGATGA 1120
Db 1003 AAAGCTGCTGATTTGCTCAGTGCAGGTCCTCAATGATGATGACATCTATTTGATGA 1062
Qy 1121 CTGCAAGATGATTTCTTATGAATTTAAAGATCTTAAATCAAGTTGATATGAGATG 1180
Db 1063 CTGCAAGATGATTTCTTATGAATTTAAAGATCTTAAATCAAGTTGATATGAGATG 1122
Qy 1181 TTTAGACTTCCAGTAACATTTTCAAGGATCAAGCCGTGTATGATATGATGATGAT 1240
Db 1123 TTTAGACTTCCAGTAACATTTTCAAGGATCAAGCCGTGTATGATATGATGATGAT 1182
Qy 1241 GTGAGAAGGGTGTCTTGGTCCATATGCGCAGGATGACCCAACTATCAATGGGGT 1300
Db 1183 GTGAGAAGGGTGTCTTGGTCCATATGCGCAGGATGACCCAACTATCAATGGGGT 1242
Qy 1301 CCTTATCAAGAAAGTCCCTATATCAAGGACAGGACCTTGTCCAGCAAAATTTGGT 1360

Db	1213	CCTTATCAAGGAAGTCCCTCATCCAGGCCAGAACTTGTCCAGCAAAACATTGGT	1302
QY	1361	GCTTTGACTCTACAAAGACCTTCCGTGATGATGTTATPACTTTGCAAGATCATCA	1420
Db	1303	GCTTTTGACTCTACAAAGACCTTCCATGATGTTATPACTTTGCAAGAGTCATCA	1362
QY	1421	GCCATGTCAATCCAGTGTTCCTATGAAACAATGCCCAATATGATCAAAACGGATGTA	1480
Db	1363	GCCATGTCAATCCAGTGTTCCTATGAAACAATGCCCAATATGATCAAAACGGATGTA	1422
QY	1481	AATTATCAATTTACAAATGTCCTGAGCCGAGTGATGCAAGATGACAGTATGAT	1540
Db	1423	AATTATCAATTTACCAAAATGTCGTAGACGAGTGATGCAAGATGCAAGATGAT	1482
QY	1541	GTTATGTTTATCGAACAAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTAAGAG	1600
Db	1483	GTTATGTTTATCGAACAAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTAAGAG	1542
QY	1601	ACTTGATA 1608	
Db	1543	ACTTGATA 1550	

```

RESULT 16
US-09-724-676-18078
; Sequence 18078, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129381.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ. ID NOS.: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18078
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-18078

```

```
Query Match      30.2%; Score 817; DB 32; Length 944;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 817; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY	193	CTGCAGCATGGGCTGGTTAACTAGAAATGTCTGTCTTTCTGGGAGATTACTTAACAGC	252
Db	9	CTGCAGCATGGGCTGGTTAACTAGAAATGTCTGTCTTTCTGGGAGATTACTTAACAGC	68
QY	253	AAGAGCAAACTATCAGAAATGGGAGAAACAATGTGCGCAAGGCTGAATAATTACTTCACAAAGA	312
Db	69	AAGAGCAAACTATCAGAAATGGGAGAAACAATGTGCGCAAGGCTGAATAATTACTTCACAAAGA	128
QY	313	AATGTTGGAATCCAACAATGTGATCACTTTCAATGGCTTTGGCCAAACGCTCCAGTTATCA	372
Db	129	AATGTTGGAATCCAACAATGTGATCACTTTCAATGGCTTTGGCCAAACGCTCCAGTTATCA	188
QY	373	TACCTTCCTTTGGATGAGAAACGAGTAGCGCTGATGTGATGTGGAGCAAAAGATCAATATT	432
Db	189	TACCTTCCTTTGGATGAGAAACGAGTAGCGCTGATGTGATGTGGAGCAAAAGATCAATATT	248
QY	433	TTCAATCGACCTGGTTAAATATCACAAGATTTTCAAAAGATGTGTGGCCAGATCTTAAAC	492
Db	249	TTCAATCGACCTGGTTAAATATCACAAGATTTTCAAAAGATGTGTGGCCAGATCTTAAAC	308
QY	493	CAGAAGAGATGAATGCACAAATGGGCTGGGAAAGACATCTGAAAGAATGTGCTAATTTCAAT	552
Db	309	CAGAAGAGATGAATGCACAAATGGGCTGGGAAAGACATCTGAAAGAATGTGCTAATTTCAAT	368
QY	553	CAAGGTAATTAAAGCATTAATCAGACATCACTGTAGCGCTGTGGAACGGGGGCTTTTCA	612
Db	369	CAAGGTAATTAAAGCATTAATCAGACATCACTGTAGCGCTGTGGAACGGGGGCTTTTCA	428

QY	613	TCGAAATTTGCACCTACATGAAATTTGGAATATCCTGAGGACAAATATTTTAAAGCTGGA	672
Db	429	TCGAAATTTGCACCTACATGAAATTTGGAATATCCTGAGGACAAATATTTTAAAGCTGGA	488
QY	673	GAACTCACATTTTGA AAAAGCGCCGTGGGAGAGATCCATATGACCCCTTAAGCTGTGACAGC	732
Db	489	GAACTCACATTTTGA AAAAGCGCCGTGGGAGAGATCCATATGACCCCTTAAGCTGTGACAGC	548
QY	723	ATCCCTTTTAATAGATGAGAAATTATCTCTGGAACTGCAGCTGATTTTATGGGGGAGAG	792
Db	549	ATCCCTTTTAATAGATGAGAAATTATCTCTGGAACTGCAGCTGATTTTATGGGGGAGAG	608
QY	793	CTTTGCTATCTTCGGAACCTTGGGGACCAACCCAAATCAGAGACAGAGCATGATTC	852
Db	609	CTTTGCTATCTTCGGAACCTTGGGGACCAACCCAAATCAGAGACAGAGCATGATTC	668
QY	853	CAGGTGGCTCATGTATCCAAAGTTGATTAGTGCCTCACTCATCTCAGAGAGTGACAAATCC	912
Db	669	CAGGTGGCTCATGTATCCAAAGTTGATTAGTGCCTCACTCATCTCAGAGAGTGACAAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCCTGGTGA AAAATGCAATATGATGAGAGAACACTCTGG	972
Db	729	TGAAGATGACAAAGTATATCTTTTCTTCCTGGTGA AAAATGCAATATGATGAGAGAACACTCTGG	788
QY	973	AAAAGCTACTACGCTAGATAGGTGCAGATATGCAAG	1009
Db	789	AAAAGCTACTACGCTAGATAGGTGCAGATATGCAAG	825

```

RESULT 17
US-09-724-676A-18078
; Sequence 18078, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18078
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-18078

```

Query Match	30.2%	Score 817;	DB 32;	Length 944;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 817; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	193	TTGCAGCATGGGCTGGTTAACTAGAGATTGTCTGTTTCTGGGGAGTTACTTACAC	252
Db	9	CTGACGATGGGCTGGTTAACTAGAGATTGTCTGTTTCTGGGGAGTTACTTACAG	68
QY	253	AAGGCAACATATCAGAAATGGGAGAACATATGCGCAGGCTGAATTATCTACAAAGA	312
Db	69	AAGGCAACATATCAGAAATGGGAGAACATATGCGCAGGCTGAATTATCTACAAAGA	128
QY	313	AATGTGGAAATCCAAACAATGTGATCACTTTCATGGCTTGCCCAACAGCTCCAGTTATCA	372
Db	129	AATGTGGAAATCCAAACAATGTGATCACTTTCATGGCTTGCCCAACAGCTCCAGTTATCA	188
QY	373	TACCTTCTTTTGGATGAGAACGAGATGAGCTGTATGTGTGGAGCAAAAGATCACAATT	432
Db	189	TACCTTCTTTTGGATGAGAACGAGATGAGCTGTATGTGTGGAGCAAAAGATCACAATT	248
QY	433	TTCAATCCAGCTGGTTAATATCAAGATTTTCAAAAAGATTGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATCCAGCTGGTTAATATCAAGATTTTCAAAAAGATTGTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAAGATGATTCAGAAAGTGGGCTGGAAAAAGACATCTGAAAGAATGTGCTAATTTTCAT	552
Db	309	CAGAAAGATGATTCAGAAAGTGGGCTGGAAAAAGACATCTGAAAGAATGTGCTAATTTTCAT	368


```
FILE REFERENCE: PB 0004 WO 5
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19786
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 0.00e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 0.00e+00
US-10-203-135-19786
```

```
Query Match 15.0%; Score 405; DB 48; Length 456;
Best Local Similarity 99.8%; Pred. No. 4.6e-172;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTCTACAA 2119
DB 1 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTCTACAA 60
QY 2120 CAGAAGATTTCAGGCAATTAACCTCTGCGATGCGGTGAACATGGTTCTTCAAACTCTT 2179
DB 61 CAGAAGATTTCAGGCAATTAACCTCTGCGATGCGGTGAACATGGTTCTTCAAACTCTT 120
QY 2180 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCAAAAGAT 2239
DB 121 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCAAAAGAT 180
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATAGATGACACTTACCCAGAG 2299
DB 181 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATAGATGACACTTACCCAGAG 240
QY 2300 GTCTGTACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATG 2359
DB 241 GTCTGTACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATG 300
QY 2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCGAAAGGCGAGACATACC 2419
DB 301 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCGAAAGGCGAGACATACC 360
QY 2420 CCAGGAGACAGTAAACAATGAAAGCACTTACAAAGAAATTAAGAAAGTGAAGAGAGG 2479
DB 361 CCAGGAGACAGTAAACAATGAAAGCACTTACAAAGAAATTAAGAAAGTGAAGAGAGG 420
QY 2480 ACCCAAGATTTGAAGAGGCGCCAGAGAGTGTCTGA 2515
DB 421 ACCCAAGATTTGAAGAGGCGCCAGAGAGTGTCTGA 456
```

RESULT 20

```
US-10-203-136-20203
Sequence 20203, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
FILE REFERENCE: PB 0004 WO 3
CURRENT APPLICATION NUMBER: US/10/203,136
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20203
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 0.00e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 0.00e+00
US-10-203-136-20203
```

```
Query Match 15.0%; Score 405; DB 48; Length 456;
Best Local Similarity 99.8%; Pred. No. 4.6e-172;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTCTACAA 2119
DB 1 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTCTACAA 60
QY 2120 CAGAAGATTTCAGGCAATTAACCTCTGCGATGCGGTGAACATGGTTCTTCAAACTCTT 2179
DB 61 CAGAAGATTTCAGGCAATTAACCTCTGCGATGCGGTGAACATGGTTCTTCAAACTCTT 120
QY 2180 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCAAAAGAT 2239
DB 121 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCAAAAGAT 180
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATAGATGACACTTACCCAGAG 2299
DB 181 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATAGATGACACTTACCCAGAG 240
QY 2300 GTCTGTACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATG 2359
DB 241 GTCTGTACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATG 300
QY 2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCGAAAGGCGAGACATACC 2419
DB 301 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCGAAAGGCGAGACATACC 360
```


Db	1	AAGTCATTGACACAGAGCATTTGGAGAAGAACTTCTTCATTAATAATGATATGAGATGGCT	60
QY	2256	CTAAGACCCAAAGAAATGTCCTAATAGCATGACACCTTAGCCAGAGGTCTGGTACAGAGACT	2315
Db	61	CTTAAGACCAAAAATAATGTCTCAATACAGACACCTTAGCCAGAGAGGTCTGGTACAGAGACT	120
QY	2316	TCATGACAGTCAATCAACCCCAATCTCAACAGCATGGATGATGTTCTGTGAACAAGTTT	2375
Db	121	TCATGACAGTCAATCAACCCCAATCTCAACAGCATGGATGATGTTCTGTGAACAAGTTT	180
QY	2376	GGAAAAGGGAACGAAACACACGTGGGCAAAAGGCCAGGACATACCCACGAGAAACAGTAACA	2435
Db	181	GGAAAAGGGAACGAAACACACGTGGGCAAAAGGCCAGGACATACCCACGAGAAACAGTAACA	240
QY	2436	AATGGAAGCATCTTACAAAGAAATTAAGAAAGGTAGAAACAGGAGGACCCACGAATTTGAGA	2495
Db	241	AATGGAAGCATCTTACAAAGAAATTAAGAAAGGTAGAAACAGGAGGACCCACGAATTTGAGA	300
QY	2496	GGGCAACCCAGAGGTCTGAGCTGCAATTAACCTCTAGAAACCTCAAAACAAGTAGAAACTTG	2555
Db	301	GGGCAACCCAGAGGTCTGAGCTGCAATTAACCTCTAGAAACCTCAAAACAAGTAGAAACTTG	360
QY	2556	CCTAGACAATAACTGGAAA	2574
Db	361	CCTAGACAATAACTGGAAA	379

```

1 RESULT 24
2 US-09-516-335-10943
3 ; Sequence 10943, Application US/09516335
4 ; GENERAL INFORMATION:
5 ; APPLICANT: Atterburn, Matthew
6 ; APPLICANT: Asghari, Vida
7 ; APPLICANT: Damavandi, Samin
8 ; APPLICANT: Dickson, Mark
9 ; APPLICANT: Drake, Jim
10 ; APPLICANT: Drmanac, Radoje
11 ; APPLICANT: Engleman, Carrie
12 ; APPLICANT: Faulkner, Brandy
13 ; APPLICANT: Fox, Melvin
14 ; APPLICANT: Garcia, Veronica
15 ; APPLICANT: Giedt, Gretchen
16 ; APPLICANT: Jesсен, Aaron
17 ; APPLICANT: Jomek, Leni
18 ; APPLICANT: Jones, Lee
19 ; APPLICANT: Kita, David
20 ; APPLICANT: Labat, Ivan
21 ; APPLICANT: Lacroix, Mimi
22 ; APPLICANT: Lomelli, Michelle
23 ; APPLICANT: Nelken, Sarah
24 ; APPLICANT: Nguyen, Kody
25 ; APPLICANT: Nguyen, Lynne
26 ; APPLICANT: Nguyen, Phuong
27 ; APPLICANT: Nogira, Margie
28 ; APPLICANT: Palencia, Servando
29 ; APPLICANT: Raisi, Fariba
30 ; APPLICANT: Randhwa, Gurpreet
31 ; APPLICANT: Sidhu, Navjivan
32 ; APPLICANT: Smith, Benjamin
33 ; APPLICANT: Smythe, Ashleigh
34 ; APPLICANT: Tkach, Joe
35 ; APPLICANT: Tran, Lien
36 ; APPLICANT: Verna, Ron
37 ; APPLICANT: Wachter, Adam
38 ; APPLICANT: Wu, James
39 ; APPLICANT: Yim, Kenneth
40 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
41 ; FILE REFERENCE: 740CIP
42 ; CURRENT APPLICATION NUMBER: US/09/516,335
43 ; EARLIER FILING DATE: 2000-03-01
44 ; EARLIER APPLICATION NUMBER: 09/321,214

```

```

; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/008,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10943
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-335-10943

```

Query Match	14.0%	Score 379;	DB 22;	Length 379;
Best Local Similarity	100.0%;	Pred. No. 2.8e-160;		
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2196	AAGTCATTGACACAGAGCATTGGAGAAACTTCTCTTAAGAAGATGATGAGATGGCT	2255	
Db	1	AAGTCATTGACACAGAGCATTGGAGAACTTCTTCTTAAGAAGATGATGAGATGGCT	60	
Qy	2256	CTAAGACCAAGAAATGTCCAATGCAATGACCACTTACCCAGAAAGTCTGGTACAGAGACT	2315	
Db	61	CTAAGACCAAGAAATGTCCAATGCAATGACCACTTACCCAGAAAGTCTGGTACAGAGACT	120	
Qy	2316	TCATGACGCTCATGAACACCCCAATCTCAACACGATGATGATGTTCTGGAAACAAATT	2375	
Db	121	TCATGACGCTCATGAACACCCCAATCTCAACACGATGATGATGTTCTGGAAACAAATT	180	

```

Db      181  GGAAGAGGAGCCGAAACCAACGTGGCGAAAGCCAGAGACCTACCCCGAGGAAAGCTAAC 240
OY      2436  AATGGAAGCACTTACAGAAATAATGAAAGGTAGAAACAGAGAGACCCAGCAATTGGAGA 2495
Db      241  AATGAAGACCTTACAGAAATAATGAAAGGTAGAAACAGAGAGACCCAGCAATTGGAGA 300
OY      2496  GGGGACCCAGGAGGTCTGAGCTGACCTTACCTCTAGAAACCTCAAAAGTAGAAACTTG 2555
Db      301  GGGGACCCAGGAGGTCTGAGCTGACCTTACCTCTAGAAACCTCAAAAGTAGAAACTTG 360
OY      2556  CCTAGACATATACCTGAAA 2574
Db      361  CCTAGACATATACCTGAAA 379

RESULT 25
US-09-733-811-10943
/ Sequence 10943, Application US/09733811
/ GENERAL INFORMATION:
/ APPLICANT: Arterburn, Matthew
/ APPLICANT: Asghari, Vida
/ APPLICANT: Damavandi, Simin
/ APPLICANT: Dickson, Mark
/ APPLICANT: Drake, Jim
/ APPLICANT: Drmanac, Radoje
/ APPLICANT: Engleman, Carrie
/ APPLICANT: Faulkner, Brandy
/ APPLICANT: Fox, Melvin
/ APPLICANT: Garcia, Veronica
/ APPLICANT: Gledet, Gretchen
/ APPLICANT: Jeesen, Aaron
/ APPLICANT: Jomek, Leni
/ APPLICANT: Jones, Lee
/ APPLICANT: Kita, David
/ APPLICANT: Labat, Ivan
/ APPLICANT: Laroya, Mimi
/ APPLICANT: Lomelli, Michelle
/ APPLICANT: Nelken, Sarah
/ APPLICANT: Nguyen, Kody
/ APPLICANT: Nguyen, Lynne
/ APPLICANT: Nguyen, Phuong
/ APPLICANT: Nogra, Margie
/ APPLICANT: Palencia, Servando
/ APPLICANT: Rastl, Farida

```

APPLICANT: Randhwa, Gurpreet
APPLICANT: Sidhu, Navjwan
APPLICANT: Smith, Benjamin
APPLICANT: Smythe, Asleigh
APPLICANT: Tkach, Joe
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 740CIP
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 09/321,214
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: 60/088,041
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ. ID NOS: 31906
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 10943
LENGTH: 379
TYPE: DNA
ORGANISM: Homo sapiens
US-09-733-811-10943

Query Match 14.0%; Score 379; DB 32; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2196 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 2255
DB 1 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 60
QY 2256 CTAAGACCAAGAAATGTCATATGATGATGATGATGATGATGATGATGAT 2315
DB 61 CTAAGACCAAGAAATGTCATATGATGATGATGATGATGATGATGATGATGAT 120
QY 2316 TCATGACGCTATCAACCAACCCCAATCTCAACAGATGATGATGATGATGATGAT 2375
DB 121 TCATGACGCTATCAACCAACCCCAATCTCAACAGATGATGATGATGATGATGAT 180
QY 2376 GGAAGAGGACCGAAGAACTGCGCAAGGCGCAAGATGATGATGATGATGATGAT 2435
DB 181 GGAAGAGGACCGAAGAACTGCGCAAGGCGCAAGATGATGATGATGATGATGAT 240
QY 2436 AATGAGAGCACTTCAAGAAATTAAGAAAGTAGAAACGAGACCCACGATTTGAGA 2495
DB 241 AATGAGAGCACTTCAAGAAATTAAGAAAGTAGAAACGAGACCCACGATTTGAGA 300
QY 2496 GGGCACCCGAGAGTGTGAGCTGATTCCTCTTGAAGAACTCAACAGATGATGATGAT 2555
DB 301 GGGCACCCGAGAGTGTGAGCTGATTCCTCTTGAAGAACTCAACAGATGATGATGAT 360
QY 2556 CCTAGACAATTAAGTGAAGAA 2574
DB 361 CCTAGACAATTAAGTGAAGAA 379

RESULT 26
US-09-733-811A-10943
Sequence 10943, Application US/09733811A

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.
APPLICANT: Drmanac, Radoje
APPLICANT: Crkvenjakov, Radomir
APPLICANT: Dickson, Mark
APPLICANT: Drmanac, Snezana
APPLICANT: Labat, Ivan
APPLICANT: Leshkowitz, Dena
APPLICANT: Kita, David
APPLICANT: Ford, John
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-740
CURRENT FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 09/321,214
PRIOR FILING DATE: 1999-05-26
NUMBER OF SEQ. ID NOS: 31906
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 10943
LENGTH: 379
TYPE: DNA
ORGANISM: Homo sapiens
US-09-733-811A-10943

Query Match 14.0%; Score 379; DB 32; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2196 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 2255
DB 1 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 60
QY 2256 CTAAGACCAAGAAATGTCATATGATGATGATGATGATGATGATGATGATGAT 2315
DB 61 CTAAGACCAAGAAATGTCATATGATGATGATGATGATGATGATGATGATGAT 120
QY 2316 TCATGACGCTATCAACCAACCCCAATCTCAACAGATGATGATGATGATGATGAT 2375
DB 121 TCATGACGCTATCAACCAACCCCAATCTCAACAGATGATGATGATGATGATGAT 180
QY 2376 GGAAGAGGACCGAAGAACTGCGCAAGGCGCAAGATGATGATGATGATGATGAT 2435
DB 181 GGAAGAGGACCGAAGAACTGCGCAAGGCGCAAGATGATGATGATGATGATGAT 240
QY 2436 AATGAGAGCACTTCAAGAAATTAAGAAAGTAGAAACGAGACCCACGATTTGAGA 2495
DB 241 AATGAGAGCACTTCAAGAAATTAAGAAAGTAGAAACGAGACCCACGATTTGAGA 300
QY 2496 GGGCACCCGAGAGTGTGAGCTGATTCCTCTTGAAGAACTCAACAGATGATGATGAT 2555
DB 301 GGGCACCCGAGAGTGTGAGCTGATTCCTCTTGAAGAACTCAACAGATGATGATGAT 360
QY 2556 CCTAGACAATTAAGTGAAGAA 2574
DB 361 CCTAGACAATTAAGTGAAGAA 379

RESULT 27

US-09-898-888-45112
Sequence 45112, Application US/09898888

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
LIBRARIES
FILE REFERENCE: 20411-748CON1
CURRENT APPLICATION NUMBER: US/09/898,888
CURRENT FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 09/340,623
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ. ID NOS: 45207
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO. 45112
LENGTH: 379
TYPE: DNA
ORGANISM: Homo sapiens
US-09-898-888-45112

Query Match 14.0%; Score 379; DB 37; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2196 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 2255
DB 1 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGGCT 60
QY 2256 CTAAGACCAAGAAATGTCATATGATGATGATGATGATGATGATGATGATGAT 2315

Db	61	CTAAGACCAAAAGAAATGTCMAATACATGACACCTAGCCAGAAAGGTCTGTATACAGACT	120
Qy	2316	TCATGCACTCATCAACGACCCCAATCTTCACACAGATGAGTAGTTCGTAAACAATTT	2375
Db	121	TCATGCACTCATCAACGACCCCAATCTTCACACAGATGAGTAGTTCGTAAACAATTT	180
Qy	2376	GGAAGAGGACCGAAGAACAGCGTCGGCAAGGCGCAGACATATCCCAAGGAACAGTAAACA	2435
Db	181	GGAAGAGGACCGAAGAACAGCGTCGGCAAGGCGCAGACATATCCCAAGGAACAGTAAACA	240
Qy	2436	AATGAAGACCTTACCAAGAAATTAAGAAAGGTAAAGAGAGGCCACCAATTTGGA	2495
Db	241	AATGAAGACCTTACCAAGAAATTAAGAAAGGTAAAGAGAGGCCACCAATTTGGA	300
Qy	2496	GGGCAACCCAGAGAGTGTCTGAGCTGATTAACCTCTAGAAACCTCAAAACAAGTAGAAACTTG	2555
Db	301	GGGCAACCCAGAGAGTGTCTGAGCTGATTAACCTCTAGAAACCTCAAAACAAGTAGAAACTTG	360
Qy	2556	CCTAGACAAATAACTGGAAA	2574
Db	361	CCTAGACAAATAACTGGAAA	379

```

RESULT 28.
US-09-898-888A-45112
Sequence 45112, Application US/09898888A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
FILE REFERENCE: 20411-748CON1
CURRENT APPLICATION NUMBER: US/09/898, 888A
CURRENT FILING DATE: 2001-07-30
PRIORITY APPLICATION NUMBER: US/09/340,623
PRIORITY FILING DATE: 1999-06-28
PRIORITY APPLICATION NUMBER: US 09/205,070
PRIORITY FILING DATE: 1998-12-03
NUMBER OF SEQ ID NOS: 45207
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 45112
LENGTH: 379
TYPE: DNA
ORGANISM: Homo sapiens
US-09-898-888A-45112

```

Query Match	14.0%;	Score 379;	DB 37;	Length 379;
Best Local Similarity	100.0%;	Pred. No. 2.8e-160;		
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	2196	AAGCATTGACACAGAGCATTTGGAAAGAACTTTCTTCAATAAAGATGATATGAGATGGCT	2255
Db	1	AAGCATTGACACAGAGCATTTGGAAAGAACTTTCTTCAATAAAGATGATGAGATGGCT	60
QY	2256	CTAAGACCCMAAGAAATGTCCAAATAGCATGACACTAGCCAGAGAGTCTGTGTAACAGAGACT	2315
Db	61	CTAAGACCCMAAGAAATGTCCAAATAGCAAGAACCTTAGCCAGAAAGGTCGTGTAACAGAGACT	120
QY	2316	TCATGACAGCTCATCAACCCACCCCAATCTCAACAGATGATGAGTTCTGTAAACAAGTTT	2375
Db	121	TCATGACAGCTCATCAACCCACCCCAATCTCAACAGATGATGAGTTCTGTAAACAAGTTT	180
QY	2376	GGAAAGAAGGACCCGAAAAACAAGTGGGCAAGGCCAGACATATCCCAAGGAAACGTATACA	2435
Db	181	GGAAAGAAGGACCCGAAAAACAAGTGGGCAAGGCCAGACATATCCCAAGGAAACGTATACA	240
QY	2436	AATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAACAGGAGGACCCACAAATTTGAGA	2495
Db	241	AATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAACAGGAGGCCACAAATTTGAGA	300
QY	2496	GGGACCCAGAGAGTCTGAGCTCATTAACCTCTTAGAAACCTCAAAACAAGTAGAAACTTG	2555
Db	301	GGGACCCAGAGAGTCTGAGCTCATTAACCTCTTAGAAACCTCAAAACAAGTAGAAACTTG	360

Qy	2556	CCTAGACAATAACTGGAAA	2574
Db	361	CCTAGACAATAACTGGAAA	379

```

RESULT 29
Sequence 10943
GENERAL INFORMATION:
  APPLICATION US/09975640
  APPLICANT: Atterburn, Matthew
  APPLICANT: Asghari, Vida
  APPLICANT: Damavandi, Simin
  APPLICANT: Dickson, Mark
  APPLICANT: Drake, Jim
  APPLICANT: Dmatnac, Radoje
  APPLICANT: Engleman, Carrie
  APPLICANT: Faulkner, Brandy
  APPLICANT: Fox, Melyin
  APPLICANT: Garcia, Veronica
  APPLICANT: Giedt, Gretchen
  APPLICANT: Uessen, Aaron
  APPLICANT: Jomek, Leni
  APPLICANT: Jones, Lee
  APPLICANT: Kita, David
  APPLICANT: Labat, Ivan
  APPLICANT: Laroya, Mimi
  APPLICANT: Lomelli, Michelle
  APPLICANT: Nelken, Sarah
  APPLICANT: Nguyen, Kody
  APPLICANT: Nguyen, Lynne
  APPLICANT: Nguyen, Phuong
  APPLICANT: Nogira, Margie
  APPLICANT: Palencia, Servando
  APPLICANT: Raissi, Farida
  APPLICANT: Randwa, Gurpreet
  APPLICANT: Sidhu, Navjwan
  APPLICANT: Smith, Benjamin
  APPLICANT: Smythe, Ashleigh
  APPLICANT: Tkach, Joe
  APPLICANT: Tran, Lien
  APPLICANT: Verna, Ron
  APPLICANT: Wachter, Adam
  APPLICANT: Wu, James
  APPLICANT: Yim, Kenneth
  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
  FILE REFERENCE: 740C1P
  CURRENT APPLICATION NUMBER: US/09/975,640
  CURRENT FILING DATE: 2001-10-11
  PRIOR APPLICATION NUMBER: 09/733,811
  PRIOR FILING DATE: 2000-12-08
  PRIOR APPLICATION NUMBER: 60/088,041
  PRIOR FILING DATE: 1998-06-02
  NUMBER OF SEQ ID NOS: 31906
  SOFTWARE: FastSeq for Windows Version 3.0
  SEQ ID NO 10943
  LENGTH: 379
  TYPE: DNA
  ORGANISM: Homo sapiens
  IS-09-975-640-10943

```

Query Match	14.0%;	Score 379;	DB 42;	Length 379;
Best Local Similarity	100.0%;	Pred. No. 2.8e-160;		
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	AAGCATTGACACAGACATTTGGAAACACTTCTTCAATAAGATGATGAATGCCT	2255
Db	1 AAGCATTTGACACAGACATTTGGAAACACTTCTTCAATAAGATGATGAATGCCT	60
QY	2256 CTAAAGCCAAAGAATATTCATATGACTGACACTAGCCAGAAGCTCTGTGACAGAGACT	2315
Db	61 CTAAAGCCAAAGAATATTCATATGACTGACACTAGCCAGAAGCTCTGTGACAGAGACT	120


```

QY 2316 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGTTCTGTGAACAAGTTT 2375
DB 121 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGTTCTGTGAACAAGTTT 180
QY 2376 GGAAGAGGACCCGAAACCAAGTCCGCAAAAGCCGAGACATACCCCAAGGAAACAGTAACA 2435
DB 181 GGAAGAGGACCCGAAACCAAGTCCGCAAAAGCCGAGACATACCCCAAGGAAACAGTAACA 240
QY 2436 AATGAGACCTTACAGAAATTAAGAAAGTAGAAACAGAGGACCCCAAGATTGAGA 2495
DB 241 AATGAGACCTTACAGAAATTAAGAAAGTAGAAACAGAGGACCCCAAGATTGAGA 300
QY 2496 GGGACCCGAGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGATGAAGACTTG 2555
DB 301 GGGACCCGAGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGATGAAGACTTG 360
QY 2556 CCTAGACATTAACCTGAAA 2574
DB 361 CCTAGACATTAACCTGAAA 379

RESULT 30
US-09-975-640A-10943
; Sequence 10943, Application US/09975640A
; GENERAL INFORMATION:
; APPLICANT: Atterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Jones, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10943
; LENGTH: 379

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640A-10943
Query Match 14.0%; Score 379; DB 42; Length 379;
Best Local Similarity 100.0%; Pred. No. 2,8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2196 AAGTCATTGACACAGAGCATTGGAAGAACTTCTTCATTAAGATGATGATGAGATGCT 2255
DB 1 AAGTCATTGACACAGAGCATTGGAAGAACTTCTTCATTAAGATGATGATGAGATGCT 60
QY 2256 CTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAGAGAGTCTGTACAGAGACT 2315
DB 61 CTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAGAGAGTCTGTACAGAGACT 120
QY 2316 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATGTTCTGTGAACAAGTTT 2375
DB 121 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATGTTCTGTGAACAAGTTT 180
QY 2376 GGAAGAGGACCCGAAACCAAGTCCGCAAAAGCCGAGACATACCCCAAGGAAACAGTAACA 2435
DB 181 GGAAGAGGACCCGAAACCAAGTCCGCAAAAGCCGAGACATACCCCAAGGAAACAGTAACA 240
QY 2436 AATGAGACCTTACAGAAATTAAGAAAGTAGAAACAGAGGACCCCAAGATTGAGA 2495
DB 241 AATGAGACCTTACAGAAATTAAGAAAGTAGAAACAGAGGACCCCAAGATTGAGA 300
QY 2496 GGGACCCGAGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGATGAAGACTTG 2555
DB 301 GGGACCCGAGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGATGAAGACTTG 360
QY 2556 CCTAGACATTAACCTGAAA 2574
DB 361 CCTAGACATTAACCTGAAA 379

RESULT 31
US-09-534-856-4
; Sequence 4, Application US/09534856
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING STRUCTURAL, SECRETED, AND
; FILE REFERENCE: PD-1015 CIP
; CURRENT APPLICATION NUMBER: US/09/534,856
; CURRENT FILING DATE: 2000-03-24
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 26334
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte ID No: hu01053939
; NAME/KEY: unsure
; LOCATION: 422
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-856-4

Query Match 13.7%; Score 370; DB 23; Length 497;
Best Local Similarity 99.8%; Pred. No. 3.3e-156;
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2163 GGTTCATACAACTCTTCTTAAGTAAACCTGGAAGTCAATGACAGAGCAATTTGAG 2222

```

Db	1	GGTTATATCAAACTCTTCTTAAAGTAACTCGTAAGTCATTGACACAGACATTTGGAAG	60
OY	2223	AACTTCTTCATTAAGATGATGTGAGATGCGTCTTAAGACCAAGAAATGTTCATTAAGCA	2282
Db	61	AACTTCTTCATTAAGATGATGTGAGATGCGTCTTAAGACCAAGAAATGTTCATTAAGCA	120
OY	2283	TGACACCTAGCCAGAAAGGCTGCTGACAGACCTTATGACGCTCATCAACCAACCCCAATC	2342
Db	121	TGACACCTAGCCAGAAAGGCTGCTGACAGACCTTATGACGCTCATCAACCAACCCCAATC	180
OY	2343	TCAACACGATGGATGAGTTCTGTGAAACAAGTTTGGAAAAAGGACCGAAAAACAACGTGGC	2402
Db	181	TCAACACATGGATGAGTTCTGTGAAACAAGTTTGGAAAAAGGACCGAAAAACAACGTGGC	240
OY	2403	AAAGGCCAGGACATTCGCCCAAGGAAACAGTAACTTAATGGAAGCATTACAGAAATTAAGA	2462
Db	241	AAAGGCCAGGACATTCGCCCAAGGAAACAGTAACTTAATGGAAGCATTACAGAAATTAAGA	300
OY	2463	AAGGTAGAAAACGAGAGGACCCACGAATTTGAGAGGGCACCCGAGAGTGTCTGAGCTGCAT	2522
Db	301	AAGGTAGAAAACGAGAGGACCCACGAATTTGAGAGGGCACCCGAGAGTGTCTGAGCTGCAT	360
OY	2523	TACCTCTAGAAAACCTCAAAACAAGTAGAATCTTGCTTAGACAATACTGAAAAAACAATG	2582
Db	361	TACCTCTAGAAAACCTCAAAACAAGTAGAATCTTGCTTAGACAATACTGAAAAAACAATG	420
OY	2583	C 2583	
Db	421	C 421	

```

RESULT 32
PCT-US02-11475A-736
; Sequence 736: Application PC/TUS0211475A
; GENERAL INFORMATION:
; APPLICANT: Corixa Corporation
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.47102PC
; CURRENT APPLICATION NUMBER: PCT/US02/11475A
; CURRENT FILING DATE: 2002-04-09
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: n = A,T,C or G
PCT-US02-11475A-736

11.2%; Score 303; DB 1; Length 354;
Best Local Similarity 99.7%; Pred. No. 7,6e-12e;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY	2098	CCTTCGTACGTGTCTACACACGAAGGATTCAGGCAATTACCTCTGCCATGGGGTGA	2157
Db	1	CTTCTGTACGTAGTCTACACACGAAGGATTCAGGCAATTACCTCTGCCATGGGNGA	60
OY	2158	ACATGGGTTCAACAACTCTTCTTAAGTAAACCTTGAAGTCATTGACACAGCATTT	2217
Db	61	ACATGGGTTCAACAACTCTTCTTAAGTAAACCTTGAAGTCATTGACACAGCATTT	120
OY	2218	GGAAGAACTTCTCATTAAGATGATGTAGAGATGGCTCTTAAGACAAAGAAATGTCCAA	2277
Db	121	GGAAGAACTTCTCATTAAGATGATGTAGAGATGGCTCTTAAGACAAAGAAATGTCCAA	180
OY	2278	TAGCATGACACTTAGCCAGAAAGTCTGTGTACAGACTTCATGCACTCATCAACACC	2337
Db	181	TAGCATGACACTTAGCCAGAAAGTCTGTGTACAGACTTCATGCACTCATCAACACC	240
OY	2338	CAATCTCAACACGATGATGATGTTCTGTGAAACAAATTGGAAAAAGGAAACGAACG	2397
Db	241	CAATCTCAACACGATGATGATGTTCTGTGAAACAAATTGGAAAAAGGAAACGAACG	300
OY	2398	TCGGCAAAAGCCAGACATACCCCAAGGAAACAGTAAACAAATGGAAGCACTTACA	2451
Db	301	TCGGCAAAAGCCAGACATACCCCAAGGAAACAGTAAACAAATGGAAGCACTTACA	354

```

RESULT 33
US-09-609-448A-736
; Sequence 736: Application US/09609448A
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secret, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C10
; CURRENT APPLICATION NUMBER: US/09/609,448A
; CURRENT FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 1081
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(354)
; OTHER INFORMATION: n = A,T,C or G
US-09-609-448A-736

Query Match      11.2%; Score 303; DB 26; Length 354;
Best Local Similarity 99.7%; Pred. No. 7,6e-126;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2098 CCTTCCTCACTGATGCTACACAGAGGATTCAGGCAATTACTCTGCGATGGGTGGA 2157
Db      1 CCTTCCTCACTGATGCTACACAGAGGATTCAGGCAATTACTCTGCGATGGGTGGA 60

QY      2158 ACATGGGTCATACAAACTCTTCTTAAAGTAAACCTTGGAAGTCATTATGACAGAGCATTT 2217
Db      61 ACATGGGTCATACAAACTCTTCTTAAAGTAAACCTTGGAAGTCATTATGACAGAGCATTT 120

QY      2218 GGAAGACTCTTTCATTAAGATGATGATGGAGTGGCTTAAGACCAAAAGAAATGTCCAA 2277
Db      121 GGAAGACTCTTTCATTAAGATGATGATGGAGTGGCTTAAGACCAAAAGAAATGTCCAA 180

QY      2278 TAGCATGACACTAGCCAGAGGCTTGTTACAGAGACTTCATGACGCTCATCAACCC 2337
Db      181 TAGCATGACACTAGCCAGAGGCTTGTTACAGAGACTTCATGACGCTCATCAACCC 240

QY      2338 CAATCTCAACGATGATGATGATTTCTGTGAACAAGTTGGAAAAAGGACCGAAAAACAAG 2397

```

Db 241 CAATCTCAACACAGATGAGTGGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 300
 Oy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 354

RESULT 34

US-09-649-811-736
 ; Sequence 736, Application US/09649811
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C11
 ; CURRENT APPLICATION NUMBER: US/09/649,811
 ; CURRENT FILING DATE: 2000-08-28
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(354)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-649-811-736

Query Match 11.2%; Score 303; DB 28; Length 354;
 Best Local Similarity 99.7%; Pred. No. 7,6e-126;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2098 CCTTGTGCTAGCTAGTCTTCAACAAGAGATTCAGGCAATTACCTCTGCGATGGGTGA 2157
 Db 1 CCTTGTGCTAGCTAGTCTTCAACAAGAGATTCAGGCAATTACCTCTGCGATGGGTGA 60
 Oy 2158 ACATGGGTTCAATACAACTCTTAAAGTAACTCTGAAAGTCAATGACAGACATTT 2217
 Db 61 ACATGGGTTCAATACAACTCTTAAAGTAACTCTGAAAGTCAATGACAGACATTT 120
 Oy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCCA 2277
 Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCCA 180
 Oy 2278 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACCAACC 2337
 Db 181 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACCAACC 240
 Oy 2338 CAATCTCAACAGATGATGAGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 2397
 Db 241 CAATCTCAACAGATGATGAGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 300
 Oy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 354

RESULT 35
 US-09-833-263-736
 ; Sequence 736, Application US/09833263
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(354)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-263-736

Query Match 11.2%; Score 303; DB 35; Length 354;
 Best Local Similarity 99.7%; Pred. No. 7,6e-126;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2098 CCTTGTGCTAGCTAGTCTTCAACAAGAGATTCAGGCAATTACCTCTGCGATGGGTGA 2157
 Db 1 CCTTGTGCTAGCTAGTCTTCAACAAGAGATTCAGGCAATTACCTCTGCGATGGGTGA 60
 Oy 2158 ACATGGGTTCAATACAACTCTTAAAGTAACTCTGAAAGTCAATGACAGACATTT 2217
 Db 61 ACATGGGTTCAATACAACTCTTAAAGTAACTCTGAAAGTCAATGACAGACATTT 120
 Oy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCCA 2277
 Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAAAGAAATGTCCA 180
 Oy 2278 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACCAACC 2337
 Db 181 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACCAACC 240
 Oy 2338 CAATCTCAACAGATGATGAGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 2397
 Db 241 CAATCTCAACAGATGATGAGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 300
 Oy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACAATGGAAGCACTTCA 354

RESULT 36
 US-09-922-217-736
 ; Sequence 736, Application US/09922217
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tonglong
 ; APPLICANT: Jiang, Yutian
 ; APPLICANT: Smith, Carole Lynn
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C13
 ; CURRENT APPLICATION NUMBER: US/09/922,217
 ; CURRENT FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 1124
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:

NAME/KEY: misc_feature
LOCATION: 57
OTHER INFORMATION: n = A,T,C or G
US-09-922-217-736

Query Match 11.2%; Score 303; DB 39; Length 354;
Best Local Similarity 99.7%; Pred. No. 7,6e-126;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2098 CCTTCGTACGAGTGTACACAGAGGATTCAGGCAATTCCTCGCATGCGGTGA 2157
DB 1 CCTTCGTACGAGTGTACACAGAGGATTCAGGCAATTCCTCGCATGCGGTGA 60
QY 2158 ACATGGGTTCAATCAAACTCTTTAAGGTAACTTGAAGTATGACAGACATTT 2217
DB 61 ACATGGGTTCAATCAAACTCTTTAAGGTAACTTGAAGTATGACAGACATTT 120
QY 2218 GGAAGAACTTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 2277
DB 121 GGAAGAACTTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 2278 TAGCATGACACCTTAAGCAGAGGCTGTGACAGACCTTATGACAGACACCC 2337
DB 181 TAGCATGACACCTTAAGCAGAGGCTGTGACAGACCTTATGACAGACACCC 240
QY 2338 CAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
DB 241 CAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 2398 TCGGCAAGGCCAGAGCATATCCCGAGGAACTTAACAAATGAGACATTACA 2451
DB 301 TCGGCAAGGCCAGAGCATATCCCGAGGAACTTAACAAATGAGACATTACA 354

RESULT 37
US-10-025-380-736
Sequence 736, Application US/10025380

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secret, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 736
LENGTH: 354
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 57
OTHER INFORMATION: n = A,T,C or G
US-10-025-380-736

Query Match 11.2%; Score 303; DB 44; Length 354;
Best Local Similarity 99.7%; Pred. No. 7,6e-126;

Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2098 CCTTCGTACGAGTGTACACAGAGGATTCAGGCAATTCCTCGCATGCGGTGA 2157
DB 1 CCTTCGTACGAGTGTACACAGAGGATTCAGGCAATTCCTCGCATGCGGTGA 60
QY 2158 ACATGGGTTCAATCAAACTCTTTAAGGTAACTTGAAGTATGACAGACATTT 2217
DB 61 ACATGGGTTCAATCAAACTCTTTAAGGTAACTTGAAGTATGACAGACATTT 120
QY 2218 GGAAGAACTTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 2277
DB 121 GGAAGAACTTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 2278 TAGCATGACACCTTAAGCAGAGGCTGTGACAGACCTTATGACAGACACCC 2337
DB 181 TAGCATGACACCTTAAGCAGAGGCTGTGACAGACCTTATGACAGACACCC 240
QY 2338 CAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
DB 241 CAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 2398 TCGGCAAGGCCAGAGCATATCCCGAGGAACTTAACAAATGAGACATTACA 2451
DB 301 TCGGCAAGGCCAGAGCATATCCCGAGGAACTTAACAAATGAGACATTACA 354

RESULT 38
US-09-864-761-1444
Sequence 1444, Application US/09864761

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408

;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; OTHER INFORMATION: MAP TO AC006322.2
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 14444
;; LENGTH: 496
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
US-09-864-761-14444

Query Match 10.8%; Score 292; DB 36; Length 496;
Best Local Similarity 100.0%; Pred. No. 7,2e-121; Indels 0; Gaps 0;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2058 AGATCAGATGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGCTAC 2117
DB 160 AGATCAGATGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGCTAC 219
QY 2118 AACAGAGGATTCAGGCAATTAACCTCTGCGCATGCGGTGAGACATGGCTTCATACAACTC 2177
DB 220 AACAGAGGATTCAGGCAATTAACCTCTGCGCATGCGGTGAGACATGGCTTCATACAACTC 279
QY 2178 TTCTTAAGTAACTCCTGGAAGTCAATTCAGACAGACATTTGGAAGAACTTCTTCAATAAG 2237
DB 280 TTCTTAAGTAACTCCTGGAAGTCAATTCAGACAGACATTTGGAAGAACTTCTTCAATAAG 339
QY 2238 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 2297
DB 340 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 399
QY 2298 AGCTCTGTACAGACACTTCATGACGCTCATCAACCCCAATCTCAAC 2349
DB 400 AGCTCTGTACAGACACTTCATGACGCTCATCAACCCCAATCTCAAC 451

RESULT 39
US-10-203-135-7281
;; Sequence 7281, Application US/10203135
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 5
;; CURRENT APPLICATION NUMBER: US/10/203,135
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7281
;; LENGTH: 496
;; TYPE: DNA

;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
US-10-203-135-7281

Query Match 10.8%; Score 292; DB 48; Length 496;
Best Local Similarity 100.0%; Pred. No. 7,2e-121; Indels 0; Gaps 0;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2058 AGATCAGATGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGCTAC 2117
DB 160 AGATCAGATGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGCTAGCTAC 219
QY 2118 AACAGAGGATTCAGGCAATTAACCTCTGCGCATGCGGTGAGACATGGCTTCATACAACTC 2177
DB 220 AACAGAGGATTCAGGCAATTAACCTCTGCGCATGCGGTGAGACATGGCTTCATACAACTC 279
QY 2178 TTCTTAAGTAACTCCTGGAAGTCAATTCAGACAGACATTTGGAAGAACTTCTTCAATAAG 2237
DB 280 TTCTTAAGTAACTCCTGGAAGTCAATTCAGACAGACATTTGGAAGAACTTCTTCAATAAG 339
QY 2238 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 2297
DB 340 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 399
QY 2298 AGCTCTGTACAGACACTTCATGACGCTCATCAACCCCAATCTCAAC 2349
DB 400 AGCTCTGTACAGACACTTCATGACGCTCATCAACCCCAATCTCAAC 451

RESULT 40
US-10-203-136-7203
;; Sequence 7203, Application US/10203136
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 3
;; CURRENT APPLICATION NUMBER: US/10/203,136
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7203
;; LENGTH: 496
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
US-10-203-136-7203

Query Match 10.8%; Score 292; DB 48; Length 496;

Best Local Similarity 100.0%; Pred. No. 7.2e-121;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2058 AGATCAGAGTGATGATCATATCATGAGACATGACAGGCTTGTCTAGAGTCTAC 2117
DB 160 AGATCAGAGTGATGATCATATCATGAGACATGACAGGCTTGTCTAGAGTCTAC 219
QY 2118 AACAGAGATGATGAGCAATTAATCTCTGCGATCGGTGGAACATGGGTATCAAACTC 2177
DB 220 AACAGAGATGATGAGCAATTAATCTCTGCGATCGGTGGAACATGGGTATCAAACTC 279
QY 2178 TTCTTAAGGTAACTCTGGAAGTATGACACAGACATTTGGAAGAACTTCTCATTAAG 2237
DB 280 TTCTTAAGGTAACTCTGGAAGTATGACACAGACATTTGGAAGAACTTCTCATTAAG 339
QY 2238 ATGATGATGAGATGGCTCTAAGACCAAGAAATGTCATATGACAGCTTACAGCCAG 2297
DB 340 ATGATGATGAGATGGCTCTAAGACCAAGAAATGTCATATGACAGCTTACAGCCAG 399
QY 2298 AGGCTGTACAGAGATCTCATGACAGCTCATCAACCAATCTCAACAC 2349
DB 400 AGGCTGTACAGAGATCTCATGACAGCTCATCAACCAATCTCAACAC 451

RESULT 41
US-09-399-720-14058
; Sequence 14058, Application US/09399720
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-777
; CURRENT APPLICATION NUMBER: US/09/399,720
; CURRENT FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 20869
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 14058
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-399-720-14058

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2454 AAAATAGAAAGGTAGAAAAGAGAGACCAAGAAATTTGAGAGGACCCAGAGAGTCT 2513
DB 102 AAAATAGAAAGGTAGAAAAGAGAGACCAAGAAATTTGAGAGGACCCAGAGAGTCT 161
QY 2514 GAGCTGATTAATCTCTAGAAAGCTCAAAAGAGTAAAGTCTGAGACATTAATGAGAA 2573
DB 162 GAGCTGATTAATCTCTAGAAAGCTCAAAAGAGTAAAGTCTGAGACATTAATGAGAA 221
QY 2574 AAACAATGCAATATACATGAATCTTTTTCATGCAATATGAGATGTTTCAATGCTG 2633
DB 222 AAACAATGCAATATACATGAATCTTTTTCATGCAATATGAGATGTTTCAATGCTG 281
QY 2634 GAAATGAGCTGAGTCCCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2693
DB 282 GAAATGAGCTGAGTCCCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 341
QY 2694 TTTTTCCTTAATACC 2709
DB 342 TTTTTCCTTAATACC 357

RESULT 42
US-09-921-378-14058
; Sequence 14058, Application US/09921378
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained

; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 20411-777
; CURRENT APPLICATION NUMBER: US/09/921,378
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/399,720
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 20869
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 14058
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-921-378-14058

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2454 AAAATAGAAAGGTAGAAAAGAGAGACCAAGAAATTTGAGAGGACCCAGAGAGTCT 2513
DB 102 AAAATAGAAAGGTAGAAAAGAGAGACCAAGAAATTTGAGAGGACCCAGAGAGTCT 161
QY 2514 GAGCTGATTAATCTCTAGAAAGCTCAAAAGAGTAAAGTCTGAGACATTAATGAGAA 2573
DB 162 GAGCTGATTAATCTCTAGAAAGCTCAAAAGAGTAAAGTCTGAGACATTAATGAGAA 221
QY 2574 AAACAATGCAATATACATGAATCTTTTTCATGCAATATGAGATGTTTCAATGCTG 2633
DB 222 AAACAATGCAATATACATGAATCTTTTTCATGCAATATGAGATGTTTCAATGCTG 281
QY 2634 GAAATGAGCTGAGTCCCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2693
DB 282 GAAATGAGCTGAGTCCCAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 341
QY 2694 TTTTTCCTTAATACC 2709
DB 342 TTTTTCCTTAATACC 357

RESULT 43
US-09-442-385-770
; Sequence 770, Application US/09442385
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Lukashov, Matvey
; TITLE OF INVENTION: Cancer Array
; FILE REFERENCE: CLON-006CIP14
; CURRENT APPLICATION NUMBER: US/09/442,385
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: 09/053,375
; NUMBER OF SEQ ID NOS: 1185
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 770
; LENGTH: 254
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic Acid Probe
US-09-442-385-770

Query Match
Best Local Similarity 100.0%; Pred. No. 1.2e-103;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1083 TCCAGGTCGAATGCAATGCACTATTTTGAATGAAGTGCAGAGATGATTTCTTAATGA 1142
DB 1 TCCAGGTCGAATGCAATGCACTATTTTGAATGAAGTGCAGAGATGATTTCTTAATGA 60
QY 1143 ACTTAAAGATCCTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1202
DB 61 ACTTAAAGATCCTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 120

QY	1203	TC AAGGAGTCA GCGCTGTGTATGTATGATCATAGATGTGAGAGGGGTTCCTTGTC	1267
Db	121	TC AAGGAGTCA GCGCTGTGTATGTATGATCATAGATGTGAGAGGGGTTCCTTGTC	180
QY	1263	CATATGCCACAGGGATGAGCCCACTATCAATGGGTGCTTATCAAGAAAGATCCCT	1322
Db	181	CATATGCCACAGGGATGAGCCCACTATCAATGGGTGCTTATCAAGAAAGATCCCT	240
QY	1323	ATCCACGGCCACGAA 1336	
Db	241	ATCCACGGCCACGAA 254	

RESULT 44
US-09-824-130-5891

Sequence 5891, Application US/09824130
GENERAL INFORMATION:
APPLICANT: Richardson, Jennifer
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THERREOR
FILE REFERENCE: 1600,2052-001
CURRENT APPLICATION NUMBER: US/09/824,130
CURRENT FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/193,420
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ. ID NOS: 8734
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5891
LENGTH: 268
TYPE: DNA
ORGANISM: Homo sapiens
US-09-824-130-5891

Query Match	8.4%;	Score 228;	DB 34;	Length 268;
Best local Similarity	100.0%;	Pred. NO. 7.3e-92;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels	

Qy	1426	GTACATCAGTGTTCCTATGAAACAATCGCCAAATAGTATCAAAACGATGTAAATTA	1485
Db	41	GTACATTCAGATGTTCCTATGAAACAATCGCCAAATAGTATCAAAACGATGTAAATTA	100
Qy	1486	TCAATTTACACAAATTTGTCTGAGACGAGTGGATCAGAAAGTGAACAGTATGATTTAT	1543
Db	101	TCAATTTACACAAATTTGTCTGAGACGAGTGGATCAGAAAGTGAACAGTATGATTTAT	160
Qy	1546	GTTTATCGGAACAGATGTGGACCGTTCCTAAAGTAGTTTCAATTCCTAAGAGACTTG	1605
Db	161	GTTTATCGGAACAGATGTGGACCGTTCCTAAAGTAGTTTCAATTCCTAAGAGACTTG	220
Qy	1606	GTATGATTTTGAAGAAGGTTCTGCTGGAAGAAATGACAGTTTTTCGGGA	1653
Db	221	GTATGATTTTGAAGAAGGTTCTGCTGGAAGAAATGACAGTTTTTCGGGA	268

RESULT 45
US-60-160-203-79/c

```

: Sequence 79, Application US/60160203
: GENERAL INFORMATION:
: APPLICANT: BONAZZI, VIVIEN
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEIN,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: C0000116

```

```

? CURRENT APPLICATION NUMBER: US/60/160,203
? CURRENT FILING DATE: 1999-10-19
? NUMBER OF SEQ ID NOS: 6374
? SOFTWARE: Fastseq for Windows Version 4.0.0
? SEQ ID NO 79
? LENGTH: 636
? TYPE: DNA
? ORGANISM: HUMAN
? US-60-160-203-79

```

Query Match	6.2%	Score 169;	DB 70;	Length 636;
Best Local Similarity	99.5%;	Pred. No. 3.8e-65;		
Matches 219; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Oy 1340 TGTCCACAAAACATTTGGTGGTTTACTCTACAAAGGACCTTCGATGATGTTATA 1399
 Db 357 TGTCACGACAAACATTTGGTGGTTTACTCTACAAAGACCTTCCTGATGATGTTATA 298

QY 1400 ACCCTTGGAGAGTCATCCAGCATGTACAATCCAGTTTCCATATGAACAATCGCCCA 1455
Db 297 ACCCTTGGAGAGTCATCCAGCATGTACAATCCAGTTTCCATATGAACAATCGCCCA 238

OY 1460 ATAGTGCATCAAAACGGATGTAATTCATTAACACAAATTGTCGTAGACCCAGTGGAT 1515
 Db 237 ATAGTGCATCAAAACGGATGTAATTCATTAACACAAATCGTCGTACACCGAGTGAT 178

QY	1520 GCAGAAATGACAGTATGATGTATTATCGGAACAG 1559
D6	177 GCGAAGATGCAGTATGATGTATTATCGGAACAG 138

RESULT 46
US-60-169-840-156/c

```

: Sequence 156, Application US/60169840
: GENERAL INFORMATION:
: APPLICANT: Bonazzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CU000164
: CURRENT APPLICATION NUMBER: US/60/169,840
: CURRENT FILING DATE: 1999-12-09
: NUMBER OF SEQ ID NOS: 9628
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 156
: LENGTH: 636
: TYPE: DNA
: ORGANISM: Human
US-60-169-840-156

```

Query Match	6.2%	Score 169;	DB 70;	Length 636;
Best Local Similarity	99.5%;	Pred. No. 3.8e-65;		
Matches 219;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500																																																																																																																											
Nb	357	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283</

QY			
1400	ACCTTTCGAAGAAGTCATCCAGCCCATGTACAATCCAGTGTTCCTATGAACAAATCGGCCA	1455	
297	ACCTTTCGAAGAAGTCATCCAGCCCATGTACAATCCAGTGTTCCTATGAACAAATCGGCCA	238	

QY 1460 ATAGTGCATCAAAACGGATGTAAATTATCAATTACACAAATTGTGTAGACCGAGTGAT 1519

237 ATAGTGCATCAAAACGGATGTAAATTATCAATTACACAAATCGTCGTGAGACCGAGTGAT 178

QY	1520	GCAGAGATGCACAGTATGATGTTATGTTATCGAACAG	1559
177	GCAGAGATGCACAGTATGATGTTATGTTATCGAACAG	138	

RESULT 47
US-10-029

```

Sequence 18184, Application US/10029386
GENERAL INFORMATION:
Applicant: Penn, Sharon G.
Applicant: Rank, David R.
Applicant: Hanzel, David K.
Title of Invention: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
Title of Invention: EXPRESSION ANALYSIS TWO
File Reference: ABOMICA-X-2
Current Application Number: US/10/029,386

```



```
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 18184
/ LENGTH: 164
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR7.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
/ OTHER INFORMATION: SWISSPROT HIT: Bf667677.1, EVALUE 9.00e-81
/ OTHER INFORMATION: NT HIT: g116162409, EVALUE 1.00e-85
US-10-029-386-18184

Query Match
Best Local Similarity 100.0%; Score 161; DB 44; Length 164;
Pred. No. 1.7e-61;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
1691 CAGCAACAATAATATATTTGTTCAAGCGCTGGGCTTCCAGCTCCCTTACACGGGTGT 1750
164 CAGCAACAATAATATATTTGTTCAAGCGCTGGGCTTCCAGCTCCCTTACACGGGTGT 105
QY 1751 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTGCTTG 1810
104 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTGCTTG 45
Db 1811 GATGTTCTGCATGTTCTGCGTATTTTCCCACTGCAGAAAG 1851
QY 44 GATGTTCTGCATGTTCTGCGTATTTTCCCACTGCAGAAAG 4

RESULT 48
US-10-029-386-4484/c
/ Sequence 4484, Application US/10029386
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Hanzel, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 4484
/ LENGTH: 591
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR7.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
/ OTHER INFORMATION: NT HIT: g116162409, EVALUE 4.00e-85
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-28
/ OTHER INFORMATION: EST_HUMAN HIT: BF667677.1, EVALUE 4.00e-80
US-10-029-386-4484

Query Match
Best Local Similarity 100.0%; Score 161; DB 44; Length 591;
Pred. No. 1.6e-61;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
1691 CAGCAACAATAATATATTTGTTCAAGCGCTGGGCTTCCAGCTCCCTTACACGGGTGT 1750
311 CAGCAACAATAATATATTTGTTCAAGCGCTGGGCTTCCAGCTCCCTTACACGGGTGT 252
QY 1751 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTGCTTG 1810
```

```
Db 251 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTGCTTG 192
QY 1811 GATGTTCTGCATGTTCTGCGTATTTTCCCACTGCAGAAAG 1851
191 GATGTTCTGCATGTTCTGCGTATTTTCCCACTGCAGAAAG 151

RESULT 49
PCT-US01-00663-20125/c
/ Sequence 20125, Application PC/TUS0100663
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Hanzel, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: PB 0004 WO 7
/ CURRENT APPLICATION NUMBER: PCT/US01/00663
/ CURRENT FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 38837
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 20125
/ LENGTH: 172
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO ACC04848.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
/ OTHER INFORMATION: NT HIT: g11421514, EVALUE 8.00e-89
/ OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
PCT-US01-00663-20125

Query Match
Best Local Similarity 100.0%; Score 160; DB 1; Length 172;
Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
310 AGAAATGTTGGAATCAACAATGTGATCATCTTCAATGCGCTTGCCCAAGCTCGAGTTA 369
160 AAAAAATGTTGGAATCAACAATGTGATCATCTTCAATGCGCTTGCCCAAGCTCGAGTTA 101
QY 370 TATATCTTCTCTTTTGATGAGAAAGAGTGTATGTGTGAGCAAGAGATCAAT 429
100 TATATCTTCTCTTTTGATGAGAAAGAGTGTATGTGTGAGCAAGAGATCAAT 41
QY 430 ATTTCATTGCACTGTTAATATATCAAGATTTTCAAG 469
40 ATTTCATTGCACTGTTAATATATCAAGATTTTCAAG 1

RESULT 50
PCT-US01-00663-20125/c
/ Sequence 20125, Application PC/TUS0100663
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
```



```
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19495
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
FEATURE: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-182-993-19495

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGGCCACAGCTCCAGTTA 369
DB 160 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGGCCACAGCTCCAGTTA 101
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACAT 429
DB 100 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACAT 41
QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 469
DB 40 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 1

RESULT 53
US-10-203-134-20052/c
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 6
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19495
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
FEATURE: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-182-993-19495
```

```
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20052
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
FEATURE: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-134-20052

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGGCCACAGCTCCAGTTA 369
DB 160 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGGCCACAGCTCCAGTTA 101
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACAT 429
DB 100 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACAT 41
QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 469
DB 40 ATTTTCATTCGACCTGTTAATATCAAGATTTCAAAAG 1

RESULT 54
US-10-203-135-19754/c
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 5
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20052
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
FEATURE: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-134-20052
```

```

; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 19754
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; FEATURE:
; OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
; US-10-203-135-19754

Query Match
Best Local Similarity 100.0%; Score 160; DB 48; Length 172;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 160 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 101

Qy 370 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
Db 100 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41

Qy 430 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 469
Db 40 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 1

RESULT 55
US-10-203-136-20171/c
; Sequence 20171, Application US/10203136
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 3
; CURRENT APPLICATION NUMBER: US/10/203,136
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38578
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 20171
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
```

```

; FEATURE:
; OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
; US-10-203-136-20171

Query Match
Best Local Similarity 100.0%; Score 160; DB 48; Length 172;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 160 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 101

Qy 370 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
Db 100 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41

Qy 430 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 469
Db 40 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 1

RESULT 56
US-10-203-137-20125/c
; Sequence 20125, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 20125
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; FEATURE:
; OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
; US-10-203-137-20125

Query Match
Best Local Similarity 100.0%; Score 160; DB 48; Length 172;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; US-10-203-137-20125
```



```
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38837
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7072
;; LENGTH: 484
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004848.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
PCT-US01-00663-7072
```

```
Query Match          5.9%; Score 160; DB 1; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 310 AGAATGTTGGAATCCAAATGTGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
    |||||||
Db 458 AGAATGTTGGAATCCAAATGTGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 399

Qy 370 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 429
    |||||||
Db 398 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 339

Qy 430 ATTTTCATTGCACTGGTTAATATCATCAAGATTTTCAAAAG 469
    |||||||
Db 338 ATTTTCATTGCACTGGTTAATATCATCAAGATTTTCAAAAG 299
```

RESULT 60

```
PCT-US01-00663-7072/c
; Sequence 7072, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 7072
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
PCT-US01-00663-7072
```

```
Query Match          5.9%; Score 160; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 310 AGAATGTTGGAATCCAAATGTGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
    |||||||
Db 458 AGAATGTTGGAATCCAAATGTGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 399

Qy 370 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 429
    |||||||
Db 398 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 339

Qy 430 ATTTTCATTGCACTGGTTAATATCATCAAGATTTTCAAAAG 469
    |||||||
Db 338 ATTTTCATTGCACTGGTTAATATCATCAAGATTTTCAAAAG 299
```

RESULT 61

```
US-09-864-761-11363/c
; Sequence 11363, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11363
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-11363

Query Match 5.9%; Score 160; DB 36; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 399
QY 370 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 299

RESULT 62
US-10-182-993-6833/C

Sequence 6833, Application US/10182993
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCES: PB 0004 WO 2
CURRENT APPLICATION NUMBER: US/10/182,993
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 6833
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-10-182-993-6833

Query Match 5.9%; Score 160; DB 47; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 399

QY 370 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 299

RESULT 63
US-10-203-134-7118/C

Sequence 7118, Application US/10203134
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCES: PB 0004 WO 6
CURRENT APPLICATION NUMBER: US/10/203,134
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7118
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
US-10-203-134-7118

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAACATGTGATCACTTCAATGCTGGCCCAAGCTCCAGTTA 399
QY 370 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAACGAGTGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGTATATATCAAGATTTTCAAAAG 299

RESULT 64
US-10-203-135-7342/C

Sequence 7342, Application US/10203135
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.


```
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
FILE REFERENCE: PB 0004 WO 5
CURRENT APPLICATION NUMBER: US/10/203,135
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37012
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7342
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-10-203-135-7342
```

```
Query Match      5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 310 AGAATGTTGGAATCCAAATGATGATCCTTCAATGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCCTTCAATGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTGGATGAGAAAGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
DB 398 TCATACCTTCTTTGGATGAGAAAGAGTAGGCTGTATGTTGGAGCAAGATCACAT 339
QY 430 ATTTTCATTGCACTGTTAATATCAAGGATTTTCAAAG 469
DB 338 ATTTTCATTGCACTGTTAATATCAAGGATTTTCAAAG 299
```

```
RESULT 65
US-10-203-136-7262/c
Sequence 7262, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
FILE REFERENCE: PB 0004 WO 3
CURRENT APPLICATION NUMBER: US/10/203,136
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
```

```
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7262
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
US-10-203-136-7262
```

```
Query Match      5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 310 AGAATGTTGGAATCCAAATGATGATCCTTCAATGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCCTTCAATGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTGGATGAGAAAGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
DB 398 TCATACCTTCTTTGGATGAGAAAGAGTAGGCTGTATGTTGGAGCAAGATCACAT 339
QY 430 ATTTTCATTGCACTGTTAATATCAAGGATTTTCAAAG 469
DB 338 ATTTTCATTGCACTGTTAATATCAAGGATTTTCAAAG 299
```

```
RESULT 66
US-10-203-137-7072/c
Sequence 7072, Application US/10203137
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: US/10/203,137
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7072
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
```

US-10-203-137-7072

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 399
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 359
QY 430 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 299

RESULT 67

US-10-203-139-7012/c

Sequence 7012, Application US/10203139

GENERAL INFORMATION:

APPLICANT: Molecular Dynamics, Inc.

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN FETAL LIVER

FILE REFERENCE: PB 0004 WO 4

CURRENT APPLICATION NUMBER: US/10/203,139

CURRENT FILING DATE: 2002-08-02

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 04 February 2000 (04.02.00)

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 26 May 2000 (26.05.00)

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 03 August 2000 (03.08.00)

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 03 October 2000 (03.10.00)

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 27 September 2000 (27.09.00)

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 21 September 2000 (21.09.00)

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 30 June 2000 (30.06.00)

NUMBER OF SEQ ID NOS: 37156

SOFTWARE: Molecular Dynamics Sequence Listing Engine

SEQ ID NO 7012

LENGTH: 484

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC004848.1

FEATURE:

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

US-10-203-139-7012

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 399
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 339
QY 430 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 469

DB 338 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 299

RESULT 68

US-10-029-386-2791

Sequence 2791, Application US/10029386

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2791

LENGTH: 526

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR7.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.68

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.69

OTHER INFORMATION: NT HIT: g15174672, EVALUE 1.00e-84

OTHER INFORMATION: EST HUMAN HIT: B1552379.1, EVALUE 2.90e+00

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 3.00e-25

US-10-029-386-2791

Query Match 5.9%; Score 160; DB 44; Length 526;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 369
DB 161 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 220
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 221 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 280
QY 430 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 469
DB 281 ATTTTCATTGACCTGCTGTTAATATCAAGATTTTCAAAAG 320

RESULT 69

US-09-864-764-11810

Sequence 11810, Application US/09864761

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 11810
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; US-09-864-761-11810

Query Match
Best Local Similarity 5.3%; Score 144; DB 36; Length 446;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GATATCACCATTGCGCAGAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTACC 1975
Db 277 GATATCACCATTGCGCAGAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTACC 336

Qy 1976 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGG 2035
Db 337 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGG 396

Qy 2036 CGAATGAAGAGCGGAAAGAGAG 2059
Db 397 CGAATGAAGAGCGGAAAGAGAG 420

RESULT 70
US-10-182-993-7713
; Sequence 7713, Application US/10182993
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 2
; ANALYSIS OF GENE EXPRESSION IN HUMAN BRAIN
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US/10/182,993
; PRIOR FILING DATE: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
```

```

; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37811
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO: 7713
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; US-10-182-993-7713

Query Match
Best Local Similarity 5.3%; Score 144; DB 47; Length 446;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GATATCACCATTGCGCAGAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTACC 1975
Db 277 GATATCACCATTGCGCAGAGCCCTGAGAGAGATCATCTATGTGTAGAGATAGTACC 336

Qy 1976 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGG 2035
Db 337 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGG 396

Qy 2036 CGAATGAAGAGCGGAAAGAGAG 2059
Db 397 CGAATGAAGAGCGGAAAGAGAG 420

RESULT 71
US-10-203-135-8417
; Sequence 8417, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 5
; ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US/10/203,135
; PRIOR FILING DATE: US 60/180,312
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO: 8417
; LENGTH: 446
; TYPE: DNA
```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
US-10-203-135-8417

Query Match
Best Local Similarity 100.0%; Score 144; DB 48; Length 446;
Pred. No. 8.3e-54; Mismatches 0; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GATATACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 1975
DB 277 GATATACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 336
QY 1976 ACATTTTGAATGCGTCGAGAGTGCAGAGAGCGCTGCTATTGGCAATTCAGAG 2035
DB 337 ACATTTTGAATGCGTCGAGAGTGCAGAGAGCGCTGCTATTGGCAATTCAGAG 396
QY 2036 CGAATGAAGCGCAAGAGAG 2059
DB 397 CGAATGAAGCGCAAGAGAG 420

RESULT 72
US-10-203-136-8331
Sequence 8331, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 3
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US/10/203,136
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/226,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 8331
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
US-10-203-136-8331

Query Match
Best Local Similarity 100.0%; Score 144; DB 48; Length 446;
Pred. No. 8.3e-54; Mismatches 0; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1916 GATATACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 1975
DB 277 GATATACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 336
QY 1976 ACATTTTGAATGCGTCGAGAGTGCAGAGAGCGCTGCTATTGGCAATTCAGAG 2035

DB 337 ACATTTTGAATGCGTCGAGAGTGCAGAGAGCGCTGCTATTGGCAATTCAGAG 396
QY 2036 CGAATGAAGCGCAAGAGAG 2059
DB 397 CGAATGAAGCGCAAGAGAG 420

RESULT 73
US-60-160-203-771/c
Sequence 771, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 771
LENGTH: 574
TYPE: DNA
ORGANISM: HUMAN
US-60-160-203-771

Query Match
Best Local Similarity 100.0%; Score 144; DB 70; Length 574;
Pred. No. 8.2e-54; Mismatches 0; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGGAAGATGACAA 925
DB 455 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGGAAGATGACAA 396
QY 926 GATATCTTTTCTCCGGAAGATGAGTGAAGAACTCTGGAAGAACTCTGAC 985
DB 395 GATATCTTTTCTCCGGAAGATGAGTGAAGAACTCTGGAAGAACTCTGAC 336
QY 986 GCTAGATAGTGCATATGCAAG 1009
DB 335 GCTAGATAGTGCATATGCAAG 312

RESULT 74
US-60-160-203-838
Sequence 838, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 838
LENGTH: 728
TYPE: DNA
ORGANISM: HUMAN
US-60-160-203-838

Query Match
Best Local Similarity 100.0%; Score 144; DB 70; Length 728;
Pred. No. 8.1e-54; Mismatches 0; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGGAAGATGACAA 925
DB 322 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGGAAGATGACAA 381
QY 926 GATATCTTTTCTCCGGAAGATGAGTGAAGAACTCTGGAAGAACTCTGAC 985

Db 382 GATACCTTTTCTCCGTAAGAAATGCAATGATGGAACACTCTGGAAGCTACTCAG 441
Qy 986 GCTAGATAGTGTAGATATGCAAG 1009
Db 442 GCTAGATAGTGTAGATATGCAAG 465

RESULT 75

US-09-864-761-28381

; Sequence 28381, Application US/09864761

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aemica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

; SEQ ID NO 28381

; LENGTH: 123

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC006322.2

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56

; OTHER INFORMATION: EST HUMAN HIT: BF667677.1, EVALUE 3.00e-63

; OTHER INFORMATION: SWISSPROT HIT: 014563, EVALUE 1.00e-18

; US-09-864-761-28381

Query Match

4.5%; Score 123; DB 36; Length 123;

Best Local Similarity 100.0%; Pred. No. 2.8e-44;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1937 COTGAAGAGAGATATCATGTGTAGAGAAATGTAGACATTTTGGAAATGCGTCCG 1996

Db 1 COTGAAGAGAGATATCATGTGTAGAGAAATGTAGACATTTTGGAAATGCGTCCG 60

Qy 1997 AAGTCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAAAATGAAGCGCAAAAGAA 2056

Db 61 AAGTCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAAAATGAAGCGCAAAAGAA 120

Qy 2057 GAG 2059

Db 121 GAG 123

Search completed: August 1, 2003, 01:56:28
Job time : 7236 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:48:41 ; Search time 327 Seconds
(without alignments)
4132.689 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatcttattatcgaatg.....aggcttttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 424677 seqs, 249437934 residues

Word size: 0

Total number of hits satisfying chosen parameters: 849354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

Pending Patents: NA.New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	1.0	1313	US-10-293-244-975	Sequence 975, App
2	23	0.8	7476	PCT-US03-09662-4	Sequence 4, Appl
3	23	0.8	7494	PCT-US03-09662-3	Sequence 3, Appl
4	23	0.8	7617	PCT-US03-09662-1	Sequence 1, Appl
5	23	0.8	7626	PCT-US03-09662-2	Sequence 2, Appl
6	21	0.8	906	US-10-293-244-459	Sequence 459, App
7	21	0.8	1026	US-10-357-930-22912	Sequence 22912, A
8	21	0.8	1026	US-10-357-930-28772	Sequence 28772, A
9	21	0.8	1181	US-10-105-837-73	Sequence 33, Appl
10	21	0.8	1538	US-10-293-244-765	Sequence 765, App
11	21	0.8	1572	US-10-105-837-664	Sequence 664, App
12	21	0.8	2028	US-10-357-930-27697	Sequence 27697, A
13	21	0.8	3091	US-09-663-481-2	Sequence 54, Appl
14	21	0.8	5511	US-10-229-541A-54	Sequence 54, Appl
15	21	0.8	5517	US-10-229-541A-79	Sequence 79, Appl
16	21	0.8	5635	US-10-229-541A-55	Sequence 55, Appl
17	21	0.8	6299	US-10-229-541A-56	Sequence 56, Appl
18	21	0.8	6299	US-10-229-541A-56	Sequence 56, Appl
19	21	0.8	6423	US-10-229-541A-57	Sequence 57, Appl
20	21	0.8	6423	US-10-229-541A-57	Sequence 57, Appl
21	20	0.7	598	US-10-357-930-50683	Sequence 50683, A
22	20	0.7	1168	US-10-105-837-552	Sequence 562, App
23	20	0.7	1423	US-10-391-363A-59	Sequence 59, Appl
24	20	0.7	1498	US-10-115-831-30	Sequence 30, Appl
25	20	0.7	1522	US-10-115-831-102	Sequence 102, App
26	20	0.7	1906	US-10-293-244-480	Sequence 480, App

27	20	0.7	1920	US-10-105-837-7	Sequence 7, Appl
28	20	0.7	2225	US-10-105-837-1046	Sequence 1046, App
29	20	0.7	2946	US-10-357-930-22213	Sequence 22213, A
30	20	0.7	2946	US-10-357-930-23051	Sequence 23051, A
31	20	0.7	2946	US-10-357-930-28070	Sequence 28070, A
32	20	0.7	2946	US-10-357-930-28916	Sequence 28916, A
33	20	0.7	3336	US-09-654-936A-134	Sequence 134, App
34	20	0.7	3336	US-10-293-244-900	Sequence 900, Appl
35	20	0.7	13831263	US-09-947-914-41	Sequence 41, Appl
36	19	0.7	433	US-10-273-573-1607	Sequence 1607, App
37	19	0.7	470	US-10-085-783A-26572	Sequence 26572, A
38	19	0.7	479	US-10-085-783A-39690	Sequence 39690, A
39	19	0.7	530	US-10-357-930-21683	Sequence 21683, A
40	19	0.7	571	US-10-357-930-27527	Sequence 27527, A
41	19	0.7	571	US-10-105-837-920	Sequence 920, App
42	19	0.7	786	US-10-273-573-4669	Sequence 4669, App
43	19	0.7	1020	US-10-357-930-22333	Sequence 22333, A
44	19	0.7	1020	US-10-357-930-24125	Sequence 24125, A
45	19	0.7	1020	US-10-357-930-29216	Sequence 29216, A
46	19	0.7	1024	US-10-105-837-866	Sequence 866, App
47	19	0.7	1099	US-10-115-831-89	Sequence 89, Appl
48	19	0.7	1102	US-10-293-244-335	Sequence 335, App
49	19	0.7	1137	US-10-357-930-24858	Sequence 24858, A
50	19	0.7	1221	US-10-105-837-162	Sequence 162, App
51	19	0.7	1244	US-10-357-930-23272	Sequence 23272, A
52	19	0.7	1244	US-10-357-930-24806	Sequence 24806, A
53	19	0.7	1244	US-10-357-930-29147	Sequence 29147, A
54	19	0.7	1252	US-10-357-930-23154	Sequence 23154, A
55	19	0.7	1252	US-10-357-930-29023	Sequence 29023, A
56	19	0.7	1281	US-10-105-837-972	Sequence 972, App
57	19	0.7	1332	US-10-105-837-196	Sequence 196, App
58	19	0.7	1354	US-10-105-837-308	Sequence 308, App
59	19	0.7	1368	US-10-105-837-871	Sequence 871, App
60	19	0.7	1370	US-10-374-780A-2359	Sequence 2359, App
61	19	0.7	1465	US-10-293-244-646	Sequence 646, App
62	19	0.7	1466	US-10-105-837-914	Sequence 914, App
63	19	0.7	1503	US-10-293-244-415	Sequence 415, App
64	19	0.7	1503	US-10-105-837-588	Sequence 588, App
65	19	0.7	1509	US-10-105-837-722	Sequence 722, App
66	19	0.7	1514	US-10-293-244-748	Sequence 748, App
67	19	0.7	1646	US-10-293-244-576	Sequence 576, App
68	19	0.7	1649	US-10-273-573-4643	Sequence 4643, App
69	19	0.7	1719	US-10-357-930-24297	Sequence 24297, A
70	19	0.7	1719	US-10-357-930-25866	Sequence 25866, A
71	19	0.7	1783	US-10-293-244-297	Sequence 297, App
72	19	0.7	1795	US-10-357-930-25092	Sequence 25092, A
73	19	0.7	1890	US-10-357-930-23084	Sequence 23084, A
74	19	0.7	1930	US-10-293-244-751	Sequence 751, App
75	19	0.7	1986	US-10-357-930-23024	Sequence 23024, A
76	19	0.7	1986	US-10-357-930-28889	Sequence 28889, A
77	19	0.7	1989	US-10-105-837-837	Sequence 837, App
78	19	0.7	2045	US-10-105-837-589	Sequence 589, App
79	19	0.7	2062	US-10-105-837-170	Sequence 170, App
80	19	0.7	2095	US-10-105-837-748	Sequence 748, App
81	19	0.7	2102	US-10-105-837-487	Sequence 487, App
82	19	0.7	2114	US-10-357-930-25185	Sequence 25185, A
83	19	0.7	2210	US-10-105-837-1009	Sequence 1009, App
84	19	0.7	2243	US-10-105-837-1008	Sequence 1008, App
85	19	0.7	2254	US-10-105-837-73	Sequence 73, Appl
86	19	0.7	2261	US-10-115-831-33	Sequence 33, Appl
87	19	0.7	2261	US-10-294-433-33	Sequence 33, Appl
88	19	0.7	2351	US-10-105-837-1007	Sequence 1007, App
89	19	0.7	2376	US-10-357-930-21284	Sequence 21284, A
90	19	0.7	2376	US-10-357-930-21872	Sequence 21872, A
91	19	0.7	2376	US-10-357-930-25228	Sequence 25228, A
92	19	0.7	2474	US-10-357-930-27126	Sequence 27126, A
93	19	0.7	2474	US-10-115-831-121	Sequence 121, App
94	19	0.7	2504	US-10-105-837-778	Sequence 778, App
95	19	0.7	2692	US-10-331-486A-50	Sequence 50, Appl
96	19	0.7	2813	US-10-357-930-30062	Sequence 30062, A
97	19	0.7	2813	US-10-357-930-30123	Sequence 30123, A
98	19	0.7	2874	US-10-105-837-10	Sequence 10, Appl
99	19	0.7	3045	US-10-105-837-9	Sequence 9, Appl

100 19 0.7 3064 6 US-10-105-837-342

Sequence 342, App

ALIGNMENTS

RESULT 1

US-10-293-244-975/c

; Sequence 975, Application US/10293244

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; APPLICANT: Tang, Y. Tom et al

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-029

; CURRENT APPLICATION NUMBER: US/10/293,244

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: Not Yet Assigned

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 09/728,422

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 09/693,325

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 09/663,561

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: 09/654,936

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 09/620,325

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/598,075

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: 09/560,875

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 09/496,914

; PRIOR FILING DATE: 2000-02-03

; NUMBER OF SEQ ID NOS: 3960

; SOFTWARE: Custom

; SEQ ID NO 975

; LENGTH: 1313

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (547)..(1239)

; US-10-293-244-975

Query Match 1.0%; Score 26; DB 6; Length 1313;

Best Local Similarity 100.0%; Pred. No. 0.00056;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTCGACCCACGCTCCGGAGTAG 80

DB 1305 GTGTCGACCCACGCTCCGGAGTAG 1280

RESULT 2

PCT-US03-09662-4/c

; Sequence 4, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Eakle, Kurt

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4

; LENGTH: 7476

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

PCT-US03-09662-4

Query Match 0.8%; Score 23; DB 1; Length 7476;

Best Local Similarity 100.0%; Pred. No. 0.03;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4495 AATCTTTATTTATCGATGTTA 4473

RESULT 3

PCT-US03-09662-3/c

; Sequence 3, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Eakle, Kurt

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 3

; LENGTH: 7494

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; PCT-US03-09662-3

Query Match 0.8%; Score 23; DB 1; Length 7494;

Best Local Similarity 100.0%; Pred. No. 0.03;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4513 AATCTTTATTTATCGATGTTA 4491

RESULT 4

PCT-US03-09662-1/c

; Sequence 1, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Eakle, Kurt

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 7617

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; PCT-US03-09662-1

Query Match 0.8%; Score 23; DB 1; Length 7617;

Best Local Similarity 100.0%; Pred. No. 0.03;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4625 AATCTTTATTTATCGATGTTA 4603

RESULT 5

PCT-US03-09662-2/c

Sequence 2, Application PC/TUS0309662
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Eakle, Kurt
APPLICANT: Imboden, Michael
TITLE OF INVENTION: Antibody Libraries
FILE REFERENCE: GALA-07772
CURRENT APPLICATION NUMBER: PCT/US03/09662
CURRENT FILING DATE: 2003-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin version 3.2
SEQ ID NO 2
LENGTH: 7626
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
PCT-US03-09662-2

Query Match 0.8%; Score 23; DB 1; Length 7626;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGAGTTA 23
|||||
DB 4634 AATCTTTATTTATCGAGTTA 4612

RESULT 6
US-10-293-244-459
Sequence 459, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 459
LENGTH: 906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (49) .. (756)
US-10-293-244-459

Query Match 0.8%; Score 21; DB 6; Length 906;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTGACCCACGCGTCCGGG 75
|||||

DB 17 GTGTGACCCACGCGTCCGGG 37

RESULT 7
US-10-357-930-22912/C
Sequence 22912, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22912
LENGTH: 1026
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1, 77, 1025, 1026
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22912

Query Match 0.8%; Score 21; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGGG 77
|||||
DB 1024 GTGTGACCCACGCGTCCGGG 1004

RESULT 8
US-10-357-930-28772/C
Sequence 28772, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314

```
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 28772
/ LENGTH: 1026
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1, 77, 1025, 1026
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28772
```

```
Query Match          0.8%; Score 21; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 1024 GTGACCCACGCGTCCGGGAG 1004
```

```
RESULT 9
US-10-105-837-33
/ Sequence 33, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pc_FL_genes Version 1.0
/ SEQ ID NO 33
/ LENGTH: 1181
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (173)..(874)
US-10-105-837-33
```

```
Query Match          0.8%; Score 21; DB 6; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 59 GTGACCCACGCGTCCGGGAG 79
```

```
RESULT 10
US-10-293-244-765
/ Sequence 765, Application US/10293244
/ GENERAL INFORMATION:
/ APPLICANT: HySeq, Inc.
/ APPLICANT: Tang, Y. Tom et al
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
US-10-105-837-664
```

```
/ FILE REFERENCE: 21272-029
/ CURRENT APPLICATION NUMBER: US/10/293,244
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: Not Yet Assigned
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 765
/ LENGTH: 1538
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (237)..(1451)
US-10-293-244-765
```

```
Query Match          0.8%; Score 21; DB 6; Length 1538;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 50 GTGACCCACGCGTCCGGGAG 70
```

```
RESULT 11
US-10-105-837-664
/ Sequence 664, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pc_FL_genes Version 1.0
/ SEQ ID NO 664
/ LENGTH: 1572
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (330)..(704)
US-10-105-837-664
```

```
Query Match          0.8%; Score 21; DB 6; Length 1572;
```

Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
|||||
Db 280 GTGACCCACGCGTCGGGAG 300

RESULT 12

US-10-357-930-27697
; Sequence 27697, Application US/10357930

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Endege, Wilson

; APPLICANT: Monahan, John

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF

; TITLE OF INVENTION: HUMAN PROSTATE CANCER

; FILE REFERENCE: MRI-007BCN

; CURRENT APPLICATION NUMBER: US/10/357,930

; PRIOR FILING DATE: 2003-02-04

; PRIOR APPLICATION NUMBER: 09/785,276

; PRIOR FILING DATE: 2003-02-16

; PRIOR APPLICATION NUMBER: 60/183,319

; PRIOR FILING DATE: 2000-02-17

; PRIOR APPLICATION NUMBER: 60/189,862

; PRIOR FILING DATE: 2000-03-16

; PRIOR APPLICATION NUMBER: 60/207,454

; PRIOR FILING DATE: 2000-05-25

; PRIOR APPLICATION NUMBER: 60/211,314

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/219,007

; PRIOR FILING DATE: 2000-07-18

; PRIOR APPLICATION NUMBER: 60/255,281

; PRIOR FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 62232

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27697

; LENGTH: 2028

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813,

; LOCATION: 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823,

; LOCATION: 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833,

; LOCATION: 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842

; OTHER INFORMATION: n = A,T,C or G

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852,

; LOCATION: 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862,

; LOCATION: 2022, 2023, 2024, 2025, 2026, 2027, 2028

; OTHER INFORMATION: n = A,T,C or G

US-10-357-930-27697

Query Match 0.8%; Score 21; DB 6; Length 2028;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
|||||
Db 7 GTGACCCACGCGTCGGGAG 27

RESULT 13

US-09-663-481-2
; Sequence 2, Application US/09663481

; GENERAL INFORMATION:

; APPLICANT: Pfizer Inc.

; APPLICANT: Fidock, Mark D.

; TITLE OF INVENTION: PHOSPHODIESTERASE ENZYMES

; FILE REFERENCE: PC10350AGPR

; CURRENT APPLICATION NUMBER: US/09/663,481

; CURRENT FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/177,326

; PRIOR FILING DATE: 2000-01-20

; PRIOR APPLICATION NUMBER: UK 9922125.1

; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 2

; LENGTH: 3091

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-663-481-2

Query Match 0.8%; Score 21; DB 5; Length 3091;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
|||||
Db 1 GTGACCCACGCGTCGGGAG 21

RESULT 14

US-10-229-541A-54
; Sequence 54, Application US/10229541A

; GENERAL INFORMATION:

; APPLICANT: McCourt, Peter

; APPLICANT: Ghaesemian, Majid

; APPLICANT: Cutler, Sean

; APPLICANT: Bonetta, Dario

; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants

; FILE REFERENCE: 22542-007CIP2

; CURRENT APPLICATION NUMBER: US/10/229,541A

; CURRENT FILING DATE: 2002-08-27

; PRIOR APPLICATION NUMBER: 10/160,764

; PRIOR FILING DATE: 2002-05-31

; PRIOR APPLICATION NUMBER: 60/294,766

; PRIOR FILING DATE: 2001-05-31

; PRIOR APPLICATION NUMBER: 60/348,909

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: 10/210,760

; PRIOR FILING DATE: 2002-08-01

; PRIOR APPLICATION NUMBER: 60/309,396

; PRIOR FILING DATE: 2001-08-01

; PRIOR APPLICATION NUMBER: 60/337,084

; PRIOR FILING DATE: 2001-12-04

; PRIOR APPLICATION NUMBER: 09/191,687

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 176

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 54

; LENGTH: 5511

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Plasmid

; OTHER INFORMATION: pBl121-35S-Ant1-AtPTB

US-10-229-541A-54

Query Match 0.8%; Score 21; DB 6; Length 5511;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
|||||
Db 3349 GTGACCCACGCGTCGGGAG 3369

RESULT 15

US-10-229-541A-79
; Sequence 79, Application US/10229541A

; GENERAL INFORMATION:

```
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ PRIOR FILING DATE: 1998-11-13
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 79
/ LENGTH: 5517
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: DN90AtFTB
/ US-10-229-541A-79
```

```
Query Match      0.8%; Score 21; DB 6; Length 5517;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCCGGAG 77
Db      3349  GTGACCCACGCGTCCGGAG 3369
```

```
RESULT 16
/ US-10-229-541A-55
/ Sequence 55, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ PRIOR FILING DATE: 1998-11-13
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 55
/ LENGTH: 5635
/ TYPE: DNA
```

```
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Plasmid
/ OTHER INFORMATION: pBI12-RD29AP-Anti-AtFTB
/ US-10-229-541A-55
```

```
Query Match      0.8%; Score 21; DB 6; Length 5635;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCCGGAG 77
Db      3473  GTGACCCACGCGTCCGGAG 3493
```

```
RESULT 17
/ US-10-229-541A-56
/ Sequence 56, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ PRIOR FILING DATE: 1998-11-13
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 56
/ LENGTH: 6299
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Plasmid
/ OTHER INFORMATION: pBI121-35S-HP-AtFTB
/ US-10-229-541A-56
```

```
Query Match      0.8%; Score 21; DB 6; Length 6299;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCCGGAG 77
Db      4883  GTGACCCACGCGTCCGGAG 4903
```

```
RESULT 18
/ US-10-229-541A-56/c
/ Sequence 56, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
```

PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,396
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 56
LENGTH: 6299
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-56

Query Match 0.8%; Score 21; DB 6; Length 6299;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACCGCGTCGGGAG 77
DB 3819 GTGACCCACCGCGTCGGGAG 3799

RESULT 19
US-10-229-541A-57
Sequence 57, Application US/10229541A
GENERAL INFORMATION:
APPLICANT: McCourt, Peter
APPLICANT: Chassemian, Majid
APPLICANT: Cutler, Sean
APPLICANT: Bonetta, Dario
TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
FILE REFERENCE: 22542-007CIP2
CURRENT APPLICATION NUMBER: US/10/229,541A
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,396
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 6423
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-57

Query Match 0.8%; Score 21; DB 6; Length 6423;
Best Local Similarity 100.0%; Pred. No. 0.41;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 57 GTGACCCACCGCGTCGGGAG 77
DB 5007 GTGACCCACCGCGTCGGGAG 5027

RESULT 20
US-10-229-541A-57/c
Sequence 57, Application US/10229541A
GENERAL INFORMATION:
APPLICANT: McCourt, Peter
APPLICANT: Chassemian, Majid
APPLICANT: Cutler, Sean
APPLICANT: Bonetta, Dario
TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
FILE REFERENCE: 22542-007CIP2
CURRENT APPLICATION NUMBER: US/10/229,541A
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 6423
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-57

Query Match 0.8%; Score 21; DB 6; Length 6423;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACCGCGTCGGGAG 77
DB 3943 GTGACCCACCGCGTCGGGAG 3923

RESULT 21
US-10-357-930-50683
Sequence 50683, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454

```
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50683
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 24
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-50683
```

```
Query Match          0.7%; Score 20; DB 6; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 744 TAGATGAGATTAATTAATCTT 763
Db 269 TAGATGAGATTAATTAATCTT 288
```

```
RESULT 22
US-10-105-837-562
; Sequence 562, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 562
; LENGTH: 1168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(615)
US-10-105-837-562
```

```
Query Match          0.7%; Score 20; DB 6; Length 1168;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGA 76
Db 25 GTGACCCACGCGTCCGGGA 44
```

```
RESULT 23
US-10-391-363A-59
; Sequence 59, Application US/10391363A
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
```

```
; APPLICANT: Valentin, Henry E
; APPLICANT: Venkatesh, Tyamagondlu V
; APPLICANT: Karunanandaa, Balasubramanian
; TITLE OF INVENTION: Homoglutathione S-transferase ("hPT") Nucleic Acids and Polypeptides
; FILE REFERENCE: REN-02-052 MON-52933
; CURRENT APPLICATION NUMBER: US/10/391,363A
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 60/365,202
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 59
; LENGTH: 1423
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-391-363A-59
```

```
Query Match          0.7%; Score 20; DB 6; Length 1423;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1372 TACAAAGACCTTCTCTGATG 1391
Db 926 TACAAAGACCTTCTCTGATG 945
```

```
RESULT 24
US-10-115-831-30
; Sequence 30, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 30
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (162)..(815)
US-10-115-831-30
```

```
Query Match          0.7%; Score 20; DB 6; Length 1498;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGA 76
Db 23 GTGACCCACGCGTCCGGGA 42
```

```
RESULT 25
US-10-115-831-102
; Sequence 102, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
```



```
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 792CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/115,831
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 102
LENGTH: 1522
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (663)..(1238)
US-10-115-831-102
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1522;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
DB 26 GTCGACCCACGCGTCGCGGA 45
```

```
RESULT 26
US-10-293-244-480
Sequence 480, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
PRIOR FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 480
LENGTH: 1906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (67)..(1740)
US-10-293-244-480
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1906;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
DB 14 GTCGACCCACGCGTCGCGGA 33
```

```
RESULT 27
US-10-105-837-7
Sequence 7, Application US/10105837
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT APPLICATION NUMBER: US/10/105,837
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 7
LENGTH: 1920
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (120)..(1394)
US-10-105-837-7
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1920;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
DB 21 GTCGACCCACGCGTCGCGGA 40
```

```
RESULT 28
US-10-105-837-1046
Sequence 1046, Application US/10105837
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT APPLICATION NUMBER: US/10/105,837
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 1046
LENGTH: 2225
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
```

```
; LOCATION: (260)...(1906)
US-10-105-837-1046

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2225;
Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 35 GTCGACCCACGCGTCCGGGA 54

RESULT 29
US-10-357-930-22213
; Sequence 22213, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22213
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2945, 2946
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22213

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2946;
Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 9 GTCGACCCACGCGTCCGGGA 28

RESULT 30
US-10-357-930-23051
; Sequence 23051, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
```

```
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23051
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2945, 2946
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23051

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2946;
Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 9 GTCGACCCACGCGTCCGGGA 28

RESULT 31
US-10-357-930-28070
; Sequence 28070, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28070
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2945, 2946
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28070
```

Query Match 0.7%; Score 20; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 9 GTGACCCACGGCTCCGGGA 28

RESULT 32

US-10-357-930-28916
; Sequence 28916, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Edege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-0078CN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28916
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2545, 2946
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28916

Query Match 0.7%; Score 20; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 9 GTGACCCACGGCTCCGGGA 28

RESULT 33

US-09-654-936A-134
; Sequence 134, Application US/09654936A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyun
; APPLICANT: Zhao, Aiqing A.
; APPLICANT: Xue, Aiqing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dimanac, Radoje T.

; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 787CIP2C
; CURRENT APPLICATION NUMBER: US/09/654,936A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 164
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 134
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1751)
US-09-654-936A-134

Query Match 0.7%; Score 20; DB 5; Length 3336;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGGCTCCGGGAG 77
|||||
Db 2 TCGACCCACGGCTCCGGGAG 21

RESULT 34
US-10-293-244-900
; Sequence 900, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 900
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1751)
US-10-293-244-900

Query Match 0.7%; Score 20; DB 6; Length 3336;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGGCTCCGGGAG 77

Db 2 TCAGACCGCGCGCGGAG 21

RESULT 35
US-09-947-914-41/C
; Sequence 41, Application US/09947914
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CL001298
; CURRENT APPLICATION NUMBER: US/09/947,914
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 75
; SEQ ID NO 41
; LENGTH: 13831263
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(13831263)
; OTHER INFORMATION: n = A,T,C or G
US-09-947-914-41

Query Match 0.7%; Score 20; DB 5; Length 13831263;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1962 TAGAGATAGTACCATTT 1981
Db 4127333 TAGAGATAGTACCATTT 4127314

RESULT 36
US-10-273-573-1607/C
; Sequence 1607, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 1607
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (362)...(252)
; OTHER INFORMATION: 97% homologous to Oryctolagus cuniculus SNRPn upstream
; OTHER INFORMATION: reading frame protein, accession number AF101043, Smith-Waterman
; OTHER INFORMATION: Score=181.
US-10-273-573-1607

Query Match 0.7%; Score 19; DB 6; Length 433;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCAAGCGCTCGGG 75
Db 429 GTCGACCAAGCGCTCGGG 411

RESULT 37
US-10-085-783A-26572
; Sequence 26572, Application US/10085783A

; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26572
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16)...(16)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-26572

Query Match 0.7%; Score 19; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2691 GGCTTTTCTCTAATACC 2709
Db 334 GGCTTTTCTCTAATACC 352

RESULT 38
US-10-085-783A-39690
; Sequence 39690, Application US/10085783A
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39690
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (356)...(356)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-39690

Query Match 0.7%; Score 19; DB 6; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1615 AGAAGAGTTCTGTGAA 1633
Db 251 AGAAGAGTTCTGTGAA 269

RESULT 39
US-10-357-930-21683


```

; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 4669
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (161)..(27)
; OTHER INFORMATION: 45% homologous to Homo sapiens type I procollagen, accession
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(786)
; OTHER INFORMATION: n = a,t,c or g
US-10-273-573-4669

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 786;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGGA 76
DB 774 TCGACCCACGCGTCCGGGA 756

RESULT 43
US-10-357-930-23333
; Sequence 23333, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23333
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1017, 1018, 1019, 1020
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23333

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1020;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGG 75
DB 6 GTCGACCCACGCGTCCGGG 24
```

```

RESULT 44
US-10-357-930-24125
; Sequence 24125, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24125
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1017, 1018, 1019, 1020
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24125

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1020;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGG 75
DB 6 GTCGACCCACGCGTCCGGG 24

RESULT 45
US-10-357-930-29216
; Sequence 29216, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
```

```

; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29216
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1017, 1018, 1019, 1020
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29216

Query Match          0.7%; Score 19; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGCGG 75
DB 6 GTGACCCACGCGTCGCGG 24

RESULT 46
US-10-105-837-866
; Sequence 866, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 866
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (519)..(980)
US-10-105-837-866

Query Match          0.7%; Score 19; DB 6; Length 1024;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGCGG 75
DB 14 GTGACCCACGCGTCGCGG 32

RESULT 47
US-10-115-831-89
; Sequence 89, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
```

```

; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792CIP2BDIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 89
; LENGTH: 1099
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140)..(445)
US-10-115-831-89

Query Match          0.7%; Score 19; DB 6; Length 1099;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGCGG 75
DB 37 GTGACCCACGCGTCGCGG 55

RESULT 48
US-10-293-244-335
; Sequence 335, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 335
; LENGTH: 1102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(408)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1102)
; OTHER INFORMATION: n = a,t,c or g
US-10-293-244-335
```


Query Match 0.7%; Score 19; DB 6; Length 1102;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 43 GTGACCCACGCGTCCGGG 61

RESULT 49

US-10-357-930-24858/c
; Sequence 24858, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24858
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 1136, 1137
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24858

Query Match 0.7%; Score 19; DB 6; Length 1137;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 1135 GTGACCCACGCGTCCGGG 1117

RESULT 50

US-10-105-837-162
; Sequence 162, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 162
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263) ..(862)
US-10-105-837-162

US-10-105-837-162

Query Match 0.7%; Score 19; DB 6; Length 1221;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 19 GTGACCCACGCGTCCGGG 37

RESULT 51

US-10-357-930-23272
; Sequence 23272, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23272
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23272

Query Match 0.7%; Score 19; DB 6; Length 1244;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 3 GTGACCCACGCGTCCGGG 21

RESULT 52


```
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ TITLE OF INVENTION: HUMAN PROSTATE CANCER
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29023
/ LENGTH: 1252
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1
/ OTHER INFORMATION: n = A,T,C or G
/ US-10-357-930-29023

Query Match          0.7%; Score 19; DB 6; Length 1252;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
         |||||
DB       1251 GTCGACCCACGCGTCGCGG 1233

RESULT 56
/ US-10-105-837-972
/ Sequence 972, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 972
/ LENGTH: 1281
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
```

```
/ LOCATION: (172) .. (1074)
/ US-10-105-837-972

Query Match          0.7%; Score 19; DB 6; Length 1281;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
         |||||
DB       14 GTCGACCCACGCGTCGCGG 32

RESULT 57
/ US-10-105-837-196
/ Sequence 196, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 196
/ LENGTH: 1332
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (171) .. (920)
/ US-10-105-837-196

Query Match          0.7%; Score 19; DB 6; Length 1332;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
         |||||
DB       40 GTCGACCCACGCGTCGCGG 58

RESULT 58
/ US-10-105-837-308
/ Sequence 308, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
```

```
SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 308
; LENGTH: 1354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (226)..(1251)
US-10-105-837-308
```

```
Query Match          0.7%: Score 19; DB 6; Length 1354;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57 GTCGACCCACGCGTCGCGG 75
      |||||
Db      129 GTCGACCCACGCGTCGCGG 147
```

```
RESULT 59
US-10-105-837-871
; Sequence 871, Application US/10105837
; GENERAL INFORMATION:
```

```
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 871
LENGTH: 1368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (103)..(447)
US-10-105-837-871
```

```
Query Match          0.7%: Score 19; DB 6; Length 1368;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57 GTCGACCCACGCGTCGCGG 75
      |||||
Db      48 GTCGACCCACGCGTCGCGG 66
```

```
RESULT 60
US-10-374-780A-2359
; Sequence 2359, Application US/10374780A
; GENERAL INFORMATION:
```

```
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creeliman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
```

```
APPLICANT: Brown, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omeira
APPLICANT: Yu, Guo-Liang
```

```
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
```

```
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2359
LENGTH: 1370
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
```

```
OTHER INFORMATION: G569
US-10-374-780A-2359
```

```
Query Match          0.7%: Score 19; DB 6; Length 1370;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57 GTCGACCCACGCGTCGCGG 75
      |||||
Db      1 GTCGACCCACGCGTCGCGG 19
```

```
RESULT 61
US-10-293-244-646
; Sequence 646, Application US/10293244
; GENERAL INFORMATION:
```

```
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
```

```
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 646
/ LENGTH: 1465
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (251)..(931)
US-10-293-244-646
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1465;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 43 GTGACCCACGCGTCCGGG 61
```

RESULT 62

```
US-10-105-837-914
/ Sequence 914, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2B.DIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 914
/ LENGTH: 1466
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (779)..(1213)
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(1466)
/ OTHER INFORMATION: n = a,t,c or g
US-10-105-837-914
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1466;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 7 GTGACCCACGCGTCCGGG 25
```

RESULT 63

```
US-10-293-244-415
/ Sequence 415, Application US/10293244
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Tang, Y. Tom et al
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 21272-029
```

```
/ CURRENT APPLICATION NUMBER: US/10/293,244
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: Not Yet Assigned
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 415
/ LENGTH: 1503
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (125)..(1474)
US-10-293-244-415
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1503;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 36 GTGACCCACGCGTCCGGG 54
```

```
RESULT 64
US-10-105-837-588
/ Sequence 588, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2B.DIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 588
/ LENGTH: 1503
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (575)..(1000)
US-10-105-837-588
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1503;
```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 40 GTGACCCACGCGTCGGG 58

RESULT 65

US-10-105-837-722
Sequence 722, Application US/10105837

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhao, Qing A.

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: Novel Nucleic Acids and

FILE REFERENCE: 784CIP2BDIVA

CURRENT FILING DATE: 2002-03-21

PRIOR APPLICATION NUMBER: 09/620,312

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1104

SOFTWARE: PL_FL_genes Version 1.0

SEQ ID NO 722

LENGTH: 1509

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (57)..(1312)

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(1509)

OTHER INFORMATION: n = a,t,c or g

US-10-105-837-722

Query Match

Best Local Similarity 0.7%; Score 19; DB 6; Length 1509;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 55 GTGACCCACGCGTCGGG 73

RESULT 66

US-10-293-244-748
Sequence 748, Application US/10293244

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

APPLICANT: Tang, Y. Tom et al

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-029

CURRENT APPLICATION NUMBER: US/10/293,244

CURRENT FILING DATE: 2002-11-12

PRIOR APPLICATION NUMBER: Not Yet Assigned

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 09/728,422

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/693,325

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/663,561

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: 09/654,936

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: 09/620,325

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/598,075

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom

SEQ ID NO 748

LENGTH: 1514

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (103)..(1374)

US-10-293-244-748

RESULT 67

US-10-293-244-576
Sequence 576, Application US/10293244

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

APPLICANT: Tang, Y. Tom et al

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-029

CURRENT APPLICATION NUMBER: US/10/293,244

CURRENT FILING DATE: 2002-11-12

PRIOR APPLICATION NUMBER: Not Yet Assigned

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 09/728,422

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/693,325

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/663,561

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: 09/654,936

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: 09/620,325

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/598,075

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom

SEQ ID NO 576

LENGTH: 1646

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (361)..(1605)

US-10-293-244-576

Query Match

Best Local Similarity 100.0%; Score 19; DB 6; Length 1646;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 22 GTGACCCACGCGTCGGG 40

```
RESULT 68
US-10-273-573-4643/C
; Sequence 4643, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 4643
; LENGTH: 1649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (737)..(1240)
; OTHER INFORMATION: 28% homologous to Drosophila melanogaster hairless
; OTHER INFORMATION: protein,accession number M95192,Smith-Waterman Score=84.
US-10-273-573-4643

Query Match
Best Local Similarity 0.7%; Score 19; DB 6; Length 1649;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTGACCCACGCGTCCG 73
Db 1630 GTGTGACCCACGCGTCCG 1612

RESULT 69
US-10-357-930-24297
; Sequence 24297, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24297
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-24297

Query Match
Best Local Similarity 0.7%; Score 19; DB 6; Length 1719;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGG 75
Db 2 GTGTGACCCACGCGTCCGG 20

RESULT 70
US-10-357-930-25866
; Sequence 25866, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25866
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25866

Query Match
Best Local Similarity 0.7%; Score 19; DB 6; Length 1719;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGG 75
Db 2 GTGTGACCCACGCGTCCGG 20

RESULT 71
US-10-293-244-297
; Sequence 297, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: NOVEL Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
```

```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGG 75
Db 2 GTGTGACCCACGCGTCCGG 20

RESULT 70
US-10-357-930-25866
; Sequence 25866, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25866
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25866

Query Match
Best Local Similarity 0.7%; Score 19; DB 6; Length 1719;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGG 75
Db 2 GTGTGACCCACGCGTCCGG 20

RESULT 71
US-10-293-244-297
; Sequence 297, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: NOVEL Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
```



```

; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 297
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78) .. (1214)
US-10-293-244-297

Query Match          0.7%; Score 19; DB 6; Length 1783;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      43 GTGACCCACGCGTCCGGG 61

RESULT 72
US-10-357-930-25092/C
; Sequence 25092, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 25092
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25092

Query Match          0.7%; Score 19; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      1790 GTGACCCACGCGTCCGGG 1772

RESULT 73
US-10-357-930-23084
; Sequence 23084, Application US/10357930
```

```

; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 23084
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-23084

Query Match          0.7%; Score 19; DB 6; Length 1890;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      9 GTGACCCACGCGTCCGGG 27

RESULT 74
US-10-293-244-751
; Sequence 751, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 751
; LENGTH: 1930
```

```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (238)..(1512)
US-10-293-244-751
```

```
Query Match 0.7%; Score 19; DB 6; Length 1930;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
Db 28 GTGACCCACGCGTCCGGG 46
```

```
RESULT 75
US-10-357-930-23024
Sequence 23024, Application US/10357930
```

```
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
PRIOR FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23024
LENGTH: 1986
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 1986
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23024
```

```
Query Match 0.7%; Score 19; DB 6; Length 1986;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
Db 10 GTGACCCACGCGTCCGGG 28
```

```
Search completed: August 1, 2003, 04:25:55
Job time : 8964 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:27:01 ; Search time 181 seconds
(without alignments)
6606.115 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatctttcttctatcgatg.....agcttttcttcttacc 2709

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfltest.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2508	92.6	2601 1	US-08-121-713D-53 Sequence 53, Appl
2	2508	92.6	2601 1	US-08-835-268-53 Sequence 53, Appl
3	2508	92.6	2601 2	US-09-060-692-53 Sequence 53, Appl
4	2508	92.6	2601 3	US-08-833-391-53 Sequence 53, Appl
5	2508	92.6	2601 4	US-09-060-610-53 Sequence 53, Appl
6	2508	92.6	2601 5	PCT-US94-10151A-53 Sequence 53, Appl
7	1253	46.3	1481 1	US-08-136-922-1 Sequence 1, Appl
8	31	1.1	7160 4	US-08-786-531B-5 Sequence 5, Appl
9	31	1.1	7235 4	US-08-786-531B-4 Sequence 4, Appl
10	22	0.8	7352 4	US-08-786-531B-1 Sequence 1, Appl
11	22	0.8	7353 4	US-08-786-531B-1 Sequence 1, Appl
12	21	0.8	1062 4	US-09-391-741A-23 Sequence 23, Appl
13	21	0.8	1062 4	US-09-391-741A-33 Sequence 33, Appl
14	21	0.8	1181 4	US-09-620-312D-33 Sequence 33, Appl
15	21	0.8	1385 3	US-08-984-288-1 Sequence 1, Appl
16	21	0.8	1572 4	US-09-620-312D-664 Sequence 664, App
17	21	0.8	1661 2	US-08-815-176-2 Sequence 2, Appl
18	21	0.8	1661 4	US-09-197-344-2 Sequence 2, Appl
19	21	0.8	2026 4	US-09-324-455-1 Sequence 1, Appl
20	20	0.7	30 3	US-08-850-961-31 Sequence 31, Appl
21	20	0.7	30 3	US-09-132-541-5 Sequence 5, Appl
22	20	0.7	30 4	US-09-479-776-31 Sequence 31, Appl
23	20	0.7	840 4	US-09-244-111-5 Sequence 5, Appl
24	20	0.7	1059 4	US-09-391-741A-9 Sequence 9, Appl
25	20	0.7	1168 4	US-09-620-312D-562 Sequence 562, App
26	20	0.7	1279 3	US-08-985-950-5 Sequence 5, Appl
27	20	0.7	1279 4	US-09-546-049-5 Sequence 5, Appl

28	20	0.7	1338 4	US-09-336-536-1 Sequence 1, Appl
29	20	0.7	1503 3	US-08-999-774A-11 Sequence 11, Appl
30	20	0.7	1772 2	US-09-482-273-96 Sequence 96, Appl
31	20	0.7	1920 4	US-09-620-312D-7 Sequence 7, Appl
32	20	0.7	2214 4	US-09-489-847-113 Sequence 113, App
33	20	0.7	2225 4	US-09-620-312D-1046 Sequence 1046, App
34	20	0.7	2227 4	US-09-489-847-30 Sequence 30, Appl
35	20	0.7	4137 3	US-09-221-235-1 Sequence 1, Appl
36	20	0.7	4137 3	US-09-221-928-1 Sequence 1, Appl
37	20	0.7	4137 3	US-09-221-527-1 Sequence 1, Appl
38	20	0.7	4137 3	US-09-221-527-1 Sequence 1, Appl
39	20	0.7	4137 3	US-09-221-236-1 Sequence 1, Appl
40	20	0.7	4137 3	US-09-221-416-1 Sequence 1, Appl
41	20	0.7	4137 3	US-09-163-115-1 Sequence 1, Appl
42	20	0.7	4137 3	US-09-163-115-1 Sequence 1, Appl
43	20	0.7	4137 3	US-09-593-553-1 Sequence 1, Appl
44	20	0.7	4137 3	US-09-221-237-1 Sequence 1, Appl
45	20	0.7	5865 3	US-08-654-737B-3 Sequence 3, Appl
46	20	0.7	6700 4	US-09-654-449-1 Sequence 1, Appl
47	20	0.7	8518 4	US-09-654-449-3 Sequence 3, Appl
48	20	0.7	8540 4	US-08-487-283A-4 Sequence 4, Appl
49	20	0.7	8540 5	PCT-US96-05611A-12 Sequence 12, Appl
50	19	0.7	256 4	US-09-313-294A-2526 Sequence 2526, App
51	19	0.7	542 3	US-09-461-697-136 Sequence 136, App
52	19	0.7	571 4	US-09-620-312D-920 Sequence 920, App
53	19	0.7	581 4	US-09-512-363-9 Sequence 9, Appl
54	19	0.7	581 4	US-09-176-200-9 Sequence 9, Appl
55	19	0.7	648 2	US-08-801-972-2 Sequence 2, Appl
56	19	0.7	648 3	US-09-178-881-2 Sequence 2, Appl
57	19	0.7	780 4	US-09-205-258-63 Sequence 63, Appl
58	19	0.7	831 4	US-09-904-615-25 Sequence 25, Appl
59	19	0.7	852 3	US-09-461-697-1 Sequence 1, Appl
60	19	0.7	9032 4	US-09-257-179-20 Sequence 20, Appl
61	19	0.7	1007 4	US-09-512-363-3 Sequence 3, Appl
62	19	0.7	1007 4	US-09-176-200-3 Sequence 3, Appl
63	19	0.7	1024 4	US-09-620-312D-866 Sequence 866, App
64	19	0.7	1037 4	US-09-489-847-112 Sequence 112, App
65	19	0.7	1116 4	US-09-372-422A-41 Sequence 41, Appl
66	19	0.7	1221 4	US-09-620-312D-162 Sequence 162, App
67	19	0.7	1281 4	US-09-620-312D-972 Sequence 972, App
68	19	0.7	1332 4	US-09-620-312D-196 Sequence 196, App
69	19	0.7	1354 4	US-09-620-312D-308 Sequence 308, App
70	19	0.7	1368 4	US-09-620-312D-871 Sequence 871, App
71	19	0.7	1390 4	US-09-205-258-124 Sequence 124, App
72	19	0.7	1411 3	US-08-964-127-5 Sequence 5, Appl
73	19	0.7	1411 4	US-09-466-692-5 Sequence 5, Appl
74	19	0.7	1411 4	US-10-000-273-5 Sequence 5, Appl
75	19	0.7	1428 3	US-09-118-442-5 Sequence 5, Appl
76	19	0.7	1428 3	US-09-677-064-5 Sequence 5, Appl
77	19	0.7	1448 4	US-09-130-491-9 Sequence 9, Appl
78	19	0.7	1466 4	US-09-620-312D-914 Sequence 914, App
79	19	0.7	1503 4	US-09-620-312D-588 Sequence 588, App
80	19	0.7	1509 4	US-09-620-312D-722 Sequence 722, App
81	19	0.7	1695 3	US-09-089-879-1 Sequence 1, Appl
82	19	0.7	1695 3	US-09-488-260-1 Sequence 1, Appl
83	19	0.7	1700 3	US-08-988-856B-3 Sequence 3, Appl
84	19	0.7	1700 5	PCT-US95-01806-3 Sequence 3, Appl
85	19	0.7	1939 4	US-09-719-083A-3 Sequence 3, Appl
86	19	0.7	1989 4	US-09-620-312D-837 Sequence 837, App
87	19	0.7	2045 4	US-09-620-312D-589 Sequence 589, App
88	19	0.7	2062 4	US-09-620-312D-1708 Sequence 1708, App
89	19	0.7	2095 4	US-09-620-312D-748 Sequence 748, App
90	19	0.7	2102 4	US-09-620-312D-487 Sequence 487, App
91	19	0.7	2210 4	US-09-620-312D-1009 Sequence 1009, App
92	19	0.7	2243 4	US-09-620-312D-1008 Sequence 1008, App
93	19	0.7	2254 4	US-09-620-312D-73 Sequence 73, Appl
94	19	0.7	2351 4	US-09-620-312D-1007 Sequence 1007, App
95	19	0.7	2504 4	US-09-620-312D-778 Sequence 778, App
96	19	0.7	2525 3	US-09-342-648-1 Sequence 1, Appl
97	19	0.7	2692 3	US-09-396-243-298 Sequence 298, App
98	19	0.7	2702 3	US-09-461-697-359 Sequence 359, App
99	19	0.7	2874 4	US-09-620-312D-10 Sequence 10, Appl
100	19	0.7	3045 4	US-09-620-312D-9 Sequence 9, Appl

|||||
Db 1329 AGTGAATCAGAGAGATGACAGTATGATATGTTATCGAAGACAGATGTTGGACCGT 1388
1573 TCTTAAAGTAGTTTCAATTCCTAAGAGAGACTTGGATATTTAGAGAGAGTTCTGCTGA 1632
1389 TCTTAAAGTAGTTTCAATTCCTAAGAGAGACTTGGATATTTAGAGAGAGTTCTGCTGA 1448
1633 AGAATGACAGTATTCGAGAACCGACTGCTATTTTCAAGCAATGAGCTTTCATTAAGCA 1692
1449 AGAATGACAGTATTCGAGAACCGACTGCTATTTTCAAGCAATGAGCTTTCATTAAGCA 1508
1693 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGTGTGA 1752
1509 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGTGTGA 1568
1753 TATTTAGGGGAAAGGCTGTCGAGTGTGCTGCGCCGAGACCTTTCTGCTGCTGGGA 1812
1569 TATTTAGGGGAAAGGCTGTCGAGTGTGCTGCGCCGAGACCTTTCTGCTGCTGGGA 1628
1813 TGGTTCTGACATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCAAGACGACAGATAT 1872
1629 TGGTTCTGACATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCAAGACGACAGATAT 1688
1873 AAGAAATGAGAGCCCACTGACTCACTGTTTCACTTACACCATGATTAATCACATGSCCA 1932
1689 AAGAAATGAGAGCCCACTGACTCACTGTTTCACTTACACCATGATTAATCACATGSCCA 1748
1933 CAGCCCTGAAAGAGAAATCATCTATNGTGTAGAGAAATAGTAGACATTTTGGATGAG 1992
1749 CAGCCCTGAAAGAGAAATCATCTATNGTGTAGAGAAATAGTAGACATTTTGGATGAG 1808
1993 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGCAATGGAAGCGAA 2052
1809 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGCAATGGAAGCGAA 1868
2053 AGAAGAGATCAGAGTGTATCATATCATGACAGACATCAAGGCTTCTGCTACGTAG 2112
1869 AGAAGAGATCAGAGTGTATCATATCATGACAGACATCAAGGCTTCTGCTACGTAG 1928
2113 TCTTCAACAGAAAGATTCAGGCAATTAACCTGCGCATGCGGTGAAACATGGGTTCAATCA 2172
1929 TCTTCAACAGAAAGATTCAGGCAATTAACCTGCGCATGCGGTGAAACATGGGTTCAATCA 1988
2173 AACTCTTCTTAAAGTAACCTCGAAGTATTTGACACAGCAATTTGGAAGAACTTCTTCA 2232
1989 AACTCTTCTTAAAGTAACCTCGAAGTATTTGACACAGCAATTTGGAAGAACTTCTTCA 2048
2233 TAAAGATGATGAGATGAGCTCTTAAAGCAAGAAATGTCATTAAGCATGACACTTAA 2292
2049 TAAAGATGATGAGATGAGCTCTTAAAGCAAGAAATGTCATTAAGCATGACACTTAA 2108
2293 CCAGAGAGTCTGATGACAGACTTCTGACAGTCTCAACCAACCCCAATCTTCAACAGAT 2352
2109 CCAGAGAGTCTGATGACAGACTTCTGACAGTCTCAACCAACCCCAATCTTCAACAGAT 2168
2353 GGATGAGTCTGATGACAGACTTCTGAGAAAGGAGCCGAAACAGCTCGGCAAGGCGCAG 2412
2169 GGATGAGTCTGATGACAGACTTCTGAGAAAGGAGCCGAAACAGCTCGGCAAGGCGCAG 2228
2413 ACATATCCCAAGGAAACATTAACAATGAGACCTTCAAGAAATAAGAAAGGTAGAA 2472
2229 ACATATCCCAAGGAAACATTAACAATGAGACCTTCAAGAAATAAGAAAGGTAGAA 2288
2473 CAGAGAGCCCAAGGAAATTTGAGAGGAGCCAGAGAGTGTGAGTGTGATTAACCTTGA 2532
2289 CAGAGAGCCCAAGGAAATTTGAGAGGAGCCAGAGAGTGTGAGTGTGATTAACCTTGA 2348
2533 AACCTCAACAGATGAGAACTTCTAGACATATAGTGAAGAAACAAATGACATATATAT 2592
2349 AACCTCAACAGATGAGAACTTCTAGACATATAGTGAAGAAACAAATGACATATATAT 2408
2593 GAACCTTTTTCATGAGCATTAATGATGATTTTCAATGATGAGAAATTCAGCTGAGTTTCA 2652
|||||

Db 2409 GAACCTTTTTCATGACATTAATGATGATTTTCAATGATGAGAAATTCAGCTGAGTTTCA 2468
Gy 2653 CCATTTTAAATTAATCCATGAGTAACCTTCTTATAGGCTTTT 2700
Db 2469 CCATTTTAAATTAATCCATGAGTAACCTTCTTATAGGCTTTT 2516
RESULT 2
US-08-835-268-53
; Sequence 53, Application US/08835268
; Patent No. 3807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mathes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-08-835-268-53
Query Match 92.6%; Score 2508; DB 1; Length 2601;
Best local similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Gy 193 CTGAGAGATGGGCTGTTAATAGATTTCTGCTTTCTGGGAGATTAATTAATTAATCAAGC 252
Db 9 CTGAGAGATGGGCTGTTAATAGATTTCTGCTTTCTGGGAGATTAATTAATTAATCAAGC 68
Gy 253 AAGAGCAAACTATCAAGATGGAAGAAACAATGTGCAAGGCTGAATTAATCTTCAAGA 312
Db 69 AAGAGCAAACTATCAAGATGGAAGAAACAATGTGCAAGGCTGAATTAATCTTCAAGA 128
Gy 313 AATGTTGAATCAACAATGATGATCACTTTCAATGCTTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTTGAATCAACAATGATGATCACTTTCAATGCTTGGCCAAAGCTCCAGTTATCA 188
|||||

OY	373	TACCTTCCTTTTGGATGAGGAACGGATGAGGCTGTATGTTGGACGAAGAATCACTATT	432
Db	189	TACCTTCCTTTTGGATGAGGAACGGATGAGGCTGTATGTTGGACGAAGAATCACTATT	248
OY	433	TTCAATCGACCTGGTATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATCGACCTGGTATATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	308
OY	493	CAGAAGATGAAATGCAAGTGGCTGGAAGAAGCATCTCGAAGAATGTGCTAATTTTCAT	552
Db	309	CAGAAGATGAAATGCAAGTGGCTGGAAGAAGCATCTCGAAGAATGTGCTAATTTTCAT	368
OY	553	CAGGATCTTAAGGCATATTAATCAGATCACTTTGATGGCTGTGGAAACGGGGCTTTTCA	612
Db	369	CAGGATCTTAAGGCATATTAATCAGATCACTTTGATGGCTGTGGAAACGGGGCTTTTCA	428
OY	613	TCCAAATTTGACCTACATTTGAATTTGACATCATCTCGTAGGACAAATTTTTTAAGCTGGA	672
Db	429	TCCAAATTTGACCTACATTTGAATTTGACATCATCTCGTAGGACAAATTTTTTAAGCTGGA	488
OY	673	GAATCACAATTTTGAAGAACGGCCGTGGGAAGATTCATATGACCCCTAGCTGTGACAC	732
Db	489	GAATCACAATTTTGAAGAACGGCCGTGGGAAGATTCATATGACCCCTAGCTGTGACAC	548
OY	723	ATCCCTTTTAATAGATGAGGAATTAATCTCTGGAATCTGACAGTATTTTAATGGGGGAGA	792
Db	549	ATCCCTTTTAATAGATGAGGAATTAATCTCTGGAATCTGACAGTATTTTAATGGGGGAGA	608
OY	793	CTTTCCTATCTTCCGAACCTTGGGGCACCAACCAATCAGGACAGACAGCATATTC	852
Db	609	CTTTCCTATCTTCCGAACCTTGGGGCACCAACCAATCAGGACAGACAGCATATTC	668
OY	853	CAGTGGCTCAATGATCCAAAGTTCATTAATGTGCCACTCATCTCAGAGATGACATTC	912
Db	669	CAGTGGCTCAATGATCCAAAGTTCATTAATGTGCCACTCATCTCAGAGATGACATTC	728
OY	913	TGAAGATGCAAGATTAATCTTTTCTCCGTGAAAATGCAATAGATGGAACATCTGTG	972
Db	729	TGAAGATGCAAGATTAATCTTTTCTCCGTGAAAATGCAATAGATGGAACATCTGTG	788
OY	973	AAAACTACTCAGCTAGAATAGTTCAGATATGCAAGAATGACTTTGAGGGCACGAAG	1032
Db	789	AAAACTACTCAGCTAGAATAGTTCAGATATGCAAGAATGACTTTGAGGGCACGAAG	848
OY	1033	TCTGTGAATTAATGACAACATTCCTCAAGCTGTGTGATTTGCTCAGTGGCCAGGTCC	1092
Db	849	TCTGTGAATTAATGACAACATTCCTCAAGCTGTGTGATTTGCTCAGTGGCCAGGTCC	908
OY	1093	AAATGGCATTGACACTCAATTTTGTAGTAACTGACAGATGTATTCCTAATGAACCTTTAANA	1152
Db	909	AAATGGCATTGACACTCAATTTTGTAGTAACTGACAGATGTATTCCTAATGAACCTTTAANA	968
OY	1153	TCTTAAAAATCCAGTTGTATATGAGAGTGTTCAGATTCCTCAGTAACATTTCAAGGATC	1212
Db	969	TCTTAAAAATCCAGTTGTATATGAGAGTGTTCAGATTCCTCAGTAACATTTCAAGGATC	1022
OY	1213	AGCCGTGTATGTATGATGATGATGTGAGAGAGTGTTCCTTGTGCTAATATGCCA	1272
Db	1029	AGCCGTGTATGTATGATGATGATGTGAGAGAGTGTTCCTTGTGCTAATATGCCA	1088
OY	1273	CAGGATGGAACCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTTATCCAGGGCC	1332
Db	1089	CAGGATGGAACCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTTATCCAGGGCC	1144
OY	1333	AGGAATCTGTCCACAGAAACATTTGGTGGTTTGACTCTACAAAGGACCTTCTGTATGA	1392
Db	1149	AGGAATCTGTCCACAGAAACATTTGGTGGTTTGACTCTACAAAGGACCTTCTGTATGA	1202
OY	1393	TGTTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACATTCAGTGTTTCTATGAACA	1452
Db	1209	TGTTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACATTCAGTGTTTCTATGAACA	1262

QY	1453	TGCCCCATATGATGCMAAAGGATGTAATTTATCAATTTACAAATTTGCTGTAGACG	1512
Db	1269	TGCCCCATATGATGCMAAAGGATGTAATTTATCAATTTACAAATTTGCTGTAGACG	1328
QY	1513	AGTGCATCAGAAAGATGAGACAGTATGATGCTATGTTTATCGGAAACAGATGTTGGGACCGT	1572
Db	1329	AGTGCATCAGAAAGATGAGACAGTATGATGCTATGTTTATCGGAAACAGATGTTGGGACCGT	1388
QY	1573	TCTTAAAGTATGTTCAATTTCTCTAAGAGACCTTGGTATGATTTAGAAAGGTTCTGTGCGA	1632
Db	1389	TCTTAAAGTATGTTCAATTTCTCTAAGAGACCTTGGTATGATTTAGAAAGGTTCTGTGCGA	1448
QY	1633	AGAAATGACAGTTTTTCCGGGAACCGACCTGCTATTTTACAGCAATTGAGAGCTTTTCCATTAACA	1692
Db	1449	AGAAATGACAGTTTTTCCGGGAACCGACCTGCTATTTTACAGCAATTGAGAGCTTTTCCATTAACA	1508
QY	1693	GCAACAACTATATATTGGTTTCAACGGGCTGGGGGTTGCCAGGCTCCCTTTACACCGGTGTGA	1752
Db	1509	GCAACAACTATATATTGGTTTCAACGGGCTGGGGGTTGCCAGGCTCCCTTTACACCGGTGTGA	1568
QY	1753	TATTTACCGGGAAGCGGTGCTGATGTTGTGCTCGCCGAGACCCCTTACTGTGCTTTGGGA	1812
Db	1569	TATTTACCGGGAAGCGGTGCTGATGTTGTGCTCGCCGAGACCCCTTACTGTGCTTTGGGA	1628
QY	1813	TGTTTCTGCATGTTCTCGCTATTTTCCACTGACAAAGACGACACAGACACAAATAT	1872
Db	1629	TGTTTCTGCATGTTCTCGCTATTTTCCACTGACAAAGACGACACAGACACAAATAT	1688
QY	1873	AAGAAATGAGAGCCCACTGACTCACTGTTTCAAGCTTACACATGATATATCAATGGCCA	1932
Db	1689	AAGAAATGAGAGCCCACTGACTCACTGTTTCAAGCTTACACATGATATATCAATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGTAACAATTTTGGAAATGACG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGTAACAATTTTGGAAATGACG	1808
QY	1993	TCCGAAGTGCAGAGAGCGCTGCTGTATTTGGCAATTCAGAGCGGAAATGAAGACGAAA	2052
Db	1809	TCCGAAGTGCAGAGAGCGCTGCTGTATTTGGCAATTCAGAGCGGAAATGAAGACGAAA	1868
QY	2053	AGAAGAGATCAGAGTGCATGATCATATCATAGGACAGATCAAGGCGCTTGCTGATCGTGA	2112
Db	1869	AGAAGAGATCAGAGTGCATGATCATATCATAGGACAGATCAAGGCGCTTGCTGATCGTGA	1928
QY	2113	TCTACAAAGAGAGATTGAGCAATTTACCTCTGCGATGCGGTGTGAAACAATGGGTTTATACA	2172
Db	1929	TCTACAAAGAGAGATTGAGCAATTTACCTCTGCGATGCGGTGTGAAACAATGGGTTTATACA	1988
QY	2173	AACTCTTTTAAAGTAACCTCTGGAAGTCAATTTGACACAGAGCATTTTGGAAAGAACTTTTCCA	2232
Db	1989	AACTCTTTTAAAGTAACCTCTGGAAGTCAATTTGACACAGAGCATTTTGGAAAGAACTTTTCCA	2048
QY	2233	TAAAGATGATATGAGAGATGGCTCTTAAGACCAAGAAATGTCAAATGACATGACACCTGAG	2292
Db	2049	TAAAGATGATATGAGAGATGGCTCTTAAGACCAAGAAATGTCAAATGACATGACACCTGAG	2108
QY	2293	CCAGAGGCTGTGTATCAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAGGCTGTGTATCAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGAAACAAGTTTGGAAAAGGACCGAAAACAAGTGGGCAAGGCCAGG	2412
Db	2169	GGATGAGTCTGTGAAACAAGTTTGGAAAAGGACCGAAAACAAGTGGGCAAGGCCAGG	2228
QY	2413	ACATATCCCAAGGAAACGTATCAAAATGGAACACTTTCAGAAATTAAGAAAGGTATGAAA	2472
Db	2229	ACATATCCCAAGGAAACGTATCAAAATGGAACACTTTCAGAAATTAAGAAAGGTATGAAA	2288
QY	2473	CAGAGAGACCAAGAAATTTGAGAGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGA	2532
Db	2289	CAGAGAGACCAAGAAATTTGAGAGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGA	2348
QY	2533	AACCTCAAAACAGTAGAAATCTTGCTGTGACATATTAACGTGAAAAACAATGCAATATACAT	2592

QY 1393 TGTATACCTTTGCAAGAGTGCATCAGCCATGTAACAATCCAGTGTTCCTATGAAACA 1452
DB 1209 TGTATACCTTTGCAAGAGTGCATCAGCCATGTAACAATCCAGTGTTCCTATGAAACA 1268
QY 1453 TGGCCCAATATGATGTAACAAAGATGTAATTTATCAATTTTACCAATTTGTGTATACCG 1512
DB 1269 TGGCCCAATATGATGTAACAAAGATGTAATTTATCAATTTTACCAATTTGTGTATACCG 1328
QY 1513 AGTGAATGCAAGATGTAAGATGATGTTATGTTATCGAAGACATGTTGGACCGT 1572
DB 1329 AGTGAATGCAAGATGTAAGATGATGTTATGTTATCGAAGACATGTTGGACCGT 1388
QY 1573 TCTTAAGTATGTTCAATTCCTTAAGAGACTTGTAATGATTTGAAGAGTTCGTCTGA 1632
DB 1389 TCTTAAGTATGTTCAATTCCTTAAGAGACTTGTAATGATTTGAAGAGTTCGTCTGA 1448
QY 1633 AGAATATGACATTTTTCGGGAACCCGACTGTATTTGAGCAATGAGACTTTCCATAGCA 1692
DB 1449 AGAATATGACATTTTTCGGGAACCCGACTGTATTTGAGCAATGAGACTTTCCATAGCA 1508
QY 1693 GCAACACTATATATGTTTCAACGAGCTGGGGTTCGCTTACACCGGTGTA 1752
DB 1509 GCAACACTATATATGTTTCAACGAGCTGGGGTTCGCTTACACCGGTGTA 1568
QY 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTTGGGA 1812
DB 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTTGGGA 1628
QY 1813 TGGTTCGATGTTTCGCTATTTTCCACTGCAAGACCGCAACAGCAAGATAT 1872
DB 1629 TGGTTCGATGTTTCGCTATTTTCCACTGCAAGACCGCAACAGCAAGATAT 1688
QY 1873 AAGAAATGAGACCCCACTGACTCACTGTTCAAGACTTACACATGATATCACATGAGCA 1932
DB 1689 AAGAAATGAGACCCCACTGACTCACTGTTCAAGACTTACACATGATATCACATGAGCA 1748
QY 1933 CAGCCCTGAAGAGAGATATCTATGTTGTAAGAAATGATGACATTTTGGAAATGCG 1992
DB 1749 CAGCCCTGAAGAGAGATATCTATGTTGTAAGAAATGATGACATTTTGGAAATGCG 1808
QY 1993 TCCGAAGTGGCAGAGAGCGCTGTCTATTTGGCAATTCGAGAGCCGAAAGAGAGCGAA 2052
DB 1809 TCCGAAGTGGCAGAGAGCGCTGTCTATTTGGCAATTCGAGAGCCGAAAGAGAGCGAA 1868
QY 2053 AGAAGAGATCAGAGTGAATCATATCATCATCAGACAGATCAAGGCTTCTGCTAGTGA 2112
DB 1869 AGAAGAGATCAGAGTGAATCATATCATCATCAGACAGATCAAGGCTTCTGCTAGTGA 1928
QY 2113 TCTACACAGAGAGATTCAGGCAATTAATCTGCGCATGCGGTGGAACATGSGTTCAATCA 2172
DB 1929 TCTACACAGAGAGATTCAGGCAATTAATCTGCGCATGCGGTGGAACATGSGTTCAATCA 1988
QY 2173 AACTCTTCTTAAGTAAACCTGGAAGTCAATGACACAGAGCAATTTGGAAAGACTTTCTTCA 2232
DB 1989 AACTCTTCTTAAGTAAACCTGGAAGTCAATGACACAGAGCAATTTGGAAAGACTTTCTTCA 2048
QY 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
DB 2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108
QY 2293 CCAAGAGTCTGTTGTAAGAGACTTCAATGAGCTCAATCAACCCCAATCTCAACAGAT 2352
DB 2109 CCAAGAGTCTGTTGTAAGAGACTTCAATGAGCTCAATCAACCCCAATCTCAACAGAT 2168
QY 2353 GATAGATGTTCTGTGAACAAATTTGAAAAAGGACCGAAACAAAGTTCGCAAGGCGCAGG 2412
DB 2169 GATAGATGTTCTGTGAACAAATTTGAAAAAGGACCGAAACAAAGTTCGCAAGGCGCAGG 2228
QY 2413 ACATACCCCAAGGAGACATGTAACAATGAAAGCACTTACAAAGAAATTAAGAAAGTGA 2472
DB 2229 ACATACCCCAAGGAGACATGTAACAATGAAAGCACTTACAAAGAAATTAAGAAAGTGA 2288

QY 2473 CAGAGAGACCAAGATTTGAGAGGGCACCAAGAGTCTGAGCTGATTAATCTCTAGA 2532
DB 2289 CAGAGAGACCAAGATTTGAGAGGGCACCAAGAGTCTGAGCTGATTAATCTCTAGA 2348
QY 2533 AACCTCAACAGATGAAACCTGCTAGACAAATTAATGAAAGAAACAAATGCAATTAAT 2592
DB 2349 AACCTCAACAGATGAAACCTGCTAGACAAATTAATGAAAGAAACAAATGCAATTAAT 2408
QY 2593 GAACTTTTTCATGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
DB 2409 GAACTTTTTCATGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
QY 2653 CCAATTAATTAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
DB 2469 CCAATTAATTAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2516

RESULT 4
US-08-833-391-53
Sequence 53, Application US/08833391
Patent No. 6013781
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-08-833-391-53

Query Match 92.6%; Score 2508; DB 3; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGAGCATGGGCTGTTAATGATGTTCTGTTTCTGGGAGATTAATTAACG 252

Db 9 CTGCGACATGGGCTGGTAACTAGATGTCTGTCCTTTCTGGGAGATTAATCTTACAGC 68
Qy 253 AAGACAAACTATCAGAAATGGAGAGACAAATGTGCGAAGGCTGAAATTAATCTTACAAAG 312
Db 69 AAGACAAACTATCAGAAATGGAGAGACAAATGTGCGAAGGCTGAAATTAATCTTACAAAG 128
Qy 313 AATGTTGAATCCAAATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGAATCCAAATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
Qy 373 TACCTTCTTTTGGATGAGAAACGAGTGGCTGATGTTGGAGCAAGAGATCAATAT 432
Db 189 TACCTTCTTTTGGATGAGAAACGAGTGGCTGATGTTGGAGCAAGAGATCAATAT 248
Qy 433 TTCAATTCGACCTGATTAATATCAAGGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 492
Db 249 TTCAATTCGACCTGATTAATATCAAGGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 308
Qy 493 CAGAAAGATGAATCAAGTGGCTGGAAGAAAGACATCCGAAAGATGTCTAATTTGAT 552
Db 309 CAGAAAGATGAATCAAGTGGCTGGAAGAAAGACATCCGAAAGATGTCTAATTTGAT 368
Qy 553 CAAGGTACTTAAGGCAATATATCAGATCTCATTGTAACGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAGGCAATATATCAGATCTCATTGTAACGCTGTGGAACGGGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTTCAATTTGAATTTGGAATCTGATCATCTGAGAGCAATTTTAACTGGA 672
Db 429 TCCAAATTTGACCTTCAATTTGAATTTGGAATCTGATCATCTGAGAGCAATTTTAACTGGA 488
Qy 673 GAACCTCACATTTTGAAGAACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
Db 489 GAACCTCACATTTTGAAGAACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
Qy 733 ATCCCTTTTAAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGCGGAGA 792
Db 549 ATCCCTTTTAAATAGATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGCGGAGA 608
Qy 793 CTTTGCTATCTTCCGAACTCTTGGGACACCAACCCATCCAGAGAGAGCATGATTC 852
Db 609 CTTTGCTATCTTCCGAACTCTTGGGACACCAACCCATCCAGAGAGAGCATGATTC 668
Qy 853 CAGGTGCTCAATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGAGTGAATCC 912
Db 669 CAGGTGCTCAATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGAGTGAATCC 728
Qy 913 TGAAGATGACAAAGTAACTTTTCTTCGTTGAAATGCAATAGATGAGAACTCTGG 972
Db 729 TGAAGATGACAAAGTAACTTTTCTTCGTTGAAATGCAATAGATGAGAACTCTGG 788
Qy 973 AAAAGCTACGACGCTAATAAATAGTCAGATATGCAATAGCAATTTGAGAGGACAGAG 1032
Db 789 AAAAGCTACGACGCTAATAAATAGTCAGATATGCAATAGCAATTTGAGAGGACAGAG 848
Qy 1033 TCTGTGATTAATAGCAACATCTCTCAAAAGCTGTCTGATTTTCTCAGTGGCCAGTCC 1092
Db 849 TCTGTGATTAATAGCAACATCTCTCAAAAGCTGTCTGATTTTCTCAGTGGCCAGTCC 908
Qy 1093 AAATGGCAATGACATCTATTTGATGAACTGCAAGATATTTCTTAATGAACTTTAAAGA 1152
Db 909 AAATGGCAATGACATCTATTTGATGAACTGCAAGATATTTCTTAATGAACTTTAAAGA 968
Qy 1153 TCCTAATAATCCAGTGTATATATGAGTGTTAACGATCCAGTAACTTTTCAAGGATC 1212
Db 969 TCCTAATAATCCAGTGTATATATGAGTGTTAACGATCCAGTAACTTTTCAAGGATC 1028
Qy 1213 AGCCGTGTATATAGCATAGATGATGAGAGGAGTGTCTTGGTCCATATAGCCCA 1272
Db 1029 AGCCGTGTATATAGCATAGATGATGAGAGGAGTGTCTTGGTCCATATAGCCCA 1088
Qy 1273 CAGGATGAGACCAATCAATCAATGGGTGCTTATAGAGAAAGTCCCTTATCCAGGCGC 1332

Db 1089 CAGGATGAGACCAATCAATGAGTGCCTTATCAAGAGAGATCCCTATCCAGGCGC 1148
Qy 1333 AGAACTGTGCCAGCAAAACATTTGGTGTGTTGACTATCAAGAGACCTTCCGTATGA 1392
Db 1149 AGAACTGTGCCAGCAAAACATTTGGTGTGTTGACTATCAAGAGACCTTCCGTATGA 1208
Qy 1393 TGTATTAACCTTTTCAAGAGATCAACAGCCATGATCAATCAGTGTTCCTATGAACA 1452
Db 1209 TGTATTAACCTTTTCAAGAGATCAACAGCCATGATCAATCAGTGTTCCTATGAACA 1268
Qy 1453 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATTAACAAATTTGTGTAGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATTAACAAATTTGTGTAGACCG 1328
Qy 1513 AGTGAATGACAGAAATGAGACAGTATGATATGTTATATGGAACAGATGTTGGAGCCGT 1572
Db 1329 AGTGAATGACAGAAATGAGACAGTATGATATGTTATATGGAACAGATGTTGGAGCCGT 1388
Qy 1573 TCTTAAAGTATGTTCAATTTCTTAAGAGACTTGTATGATTTAGAAAGAGTTCTGTGGA 1632
Db 1389 TCTTAAAGTATGTTCAATTTCTTAAGAGACTTGTATGATTTAGAAAGAGTTCTGTGGA 1448
Qy 1633 AGAAATGACAGTTTTGGGAAACCGACTGCTATTTACGAATGAGAGCTTTCCACTAAGCA 1692
Db 1449 AGAAATGACAGTTTTGGGAAACCGACTGCTATTTACGAATGAGAGCTTTCCACTAAGCA 1508
Qy 1693 GCACCACTATATATGTTTCAACGCGTGGGTTGCCAGTCCCTTTACACCGGTGGA 1752
Db 1509 GCACCACTATATATGTTTCAACGCGTGGGTTGCCAGTCCCTTTACACCGGTGGA 1568
Qy 1753 TATTTAGGGGAAACGTTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTCTGGGA 1812
Db 1569 TATTTAGGGGAAACGTTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGTCTGGGA 1628
Qy 1813 TGTGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGAGCAAGATAT 1872
Db 1629 TGTGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGAGCAAGATAT 1688
Qy 1873 AAGAAATGAGAGACCACTGACTCATCTGTGAGATTCACCATGATATACCAATGGCCA 1992
Db 1689 AAGAAATGAGAGACCACTGACTCATCTGTGAGATTCACCATGATATACCAATGGCCA 1748
Qy 1933 CAGCCTGAGAGAGAAATCAATCTATGTTGATAGAAATAGTAGACATTTTGTGAATCAG 1992
Db 1749 CAGCCTGAGAGAGAAATCAATCTATGTTGATAGAAATAGTAGACATTTTGTGAATCAG 1808
Qy 1993 TCCGAATGCGAGAGACGCTGTCTATTTGGCAATTCAGAGCGGAATGAGAGCGAAA 2052
Db 1809 TCCGAATGCGAGAGACGCTGTCTATTTGGCAATTCAGAGCGGAATGAGAGCGAAA 1868
Qy 2053 AGAAGATCAGAGTGTGATATCAATCAATCAGAGACATCAAGGCTTCTGCTACGTAG 2112
Db 1869 AGAAGATCAGAGTGTGATATCAATCAATCAGAGACATCAAGGCTTCTGCTACGTAG 1928
Qy 2113 TCTTCAACAGAGAGATTCAGGCAATTAATCTCTGCAAGCGGTGGAATGAGGTTTCATACA 2172
Db 1929 TCTTCAACAGAGAGATTCAGGCAATTAATCTCTGCAAGCGGTGGAATGAGGTTTCATACA 1968
Qy 2173 AACTCTTCTTAAGGTAACCTTGAAGTCAATGACACAGACATTTGGAAGAACTTCTTCA 2232
Db 1969 AACTCTTCTTAAGGTAACCTTGAAGTCAATGACACAGACATTTGGAAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGATGAGATGAGTCTTAAGACCAAGAAATGTCAAATGATGACACTAG 2252
Db 2049 TAAAGATGATGATGAGATGAGTCTTAAGACCAAGAAATGTCAAATGATGACACTAG 2108
Qy 2293 CCAAGAGTGTGATACAGACATTCAGTCAATCAACACCCATCTCAACAGAT 2352
Db 2109 CCAAGAGTGTGATACAGACATTCAGTCAATCAACACCCATCTCAACAGAT 2168
Qy 2353 GGATGAGTGTCTGTAACAGTGTGGAAGAGGACCGAAACAACTGTGGCAAGGCGCAGG 2412
Db 2169 GGATGAGTGTCTGTAACAGTGTGGAAGAGGACCGAAACAACTGTGGCAAGGCGCAGG 2228


```

Db      1029  AGCCCTGTGTATGTATAGCATAGTGAATGTGAAGAGGGGTCTTGGTCCATATAGCCCA 1088
Oy      1273  CAGGGATGGACCCACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCAGGCGC 1332
Db      1089  CAGGGATGGACCCACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCAGGCGC 1148
Oy      1333  AGGAACCTTGCCAGCAAAACATTTGGTGGTTTGAAGCTCTCAAGAAAGACCTTCTGATGA 1392
Db      1149  AGGAACCTTGCCAGCAAAACATTTGGTGGTTTGAAGCTCTCAAGAAAGACCTTCTGATGA 1208
Oy      1393  TGTATTAACCTTTGCAAGAGATCATTCAGCCATGTACAAATCCAGTGTTCCTATGAACA 1452
Db      1209  TGTATTAACCTTTGCAAGAGATCATTCAGCCATGTACAAATCCAGTGTTCCTATGAACA 1268
Oy      1453  TCGCCCATATGATCAAAAGGAGTAAATATATCAATTTTCAAAATTTGCTGTAGACG 1512
Db      1289  TCGCCCATATGATCAAAAGGAGTAAATATATCAATTTTCAAAATTTGCTGTAGACG 1328
Oy      1513  AGTGATCGAGAAATGAGACAGTATGATGTTATGTTATCGGAACAGATGTTGGAGCCGT 1572
Db      1329  AGTGATCGAGAAATGAGACAGTATGATGTTATGTTATCGGAACAGATGTTGGAGCCGT 1388
Oy      1573  TCTTAAAGTATGTTCAATTTCTTAAAGAGACCTTGGTATGATTTGAAGAGGTTCTGTGGA 1632
Db      1389  TCTTAAAGTATGTTCAATTTCTTAAAGAGACCTTGGTATGATTTGAAGAGGTTCTGTGGA 1448
Oy      1633  AGAAATGACAGTGTTCGGGGAACCGACGTCTATTTCAACAATGAGCTTCCACATAGCA 1692
Db      1449  AGAAATGACAGTGTTCGGGGAACCGACGTCTATTTCAACAATGAGCTTCCACATAGCA 1508
Oy      1693  GCAACAACTATATATATGTTTCAACGCGCTGGGTTCCAGACTCCCTTTACACCGGTGGA 1752
Db      1509  GCAACAACTATATATATGTTTCAACGCGCTGGGTTCCAGACTCCCTTTACACCGGTGGA 1568
Oy      1753  TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGAACCTTACCTGTGCTTGGGA 1812
Db      1569  TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGAACCTTACCTGTGCTTGGGA 1628
Oy      1813  TGGTTCGTCATGTTCTGCTATTTTCCAGTGGCAAGAGAGGCAACAAAGCAAGATAT 1872
Db      1629  TGGTTCGTCATGTTCTGCTATTTTCCAGTGGCAAGAGAGGCAACAAAGCAAGATAT 1688
Oy      1873  AAGAATATGAGACCCACTGACTGACTGTTAGACTTACCAACATGATATCAACATGGCCA 1932
Db      1689  AAGAATATGAGACCCACTGACTGACTGTTAGACTTACCAACATGATATCAACATGGCCA 1748
Oy      1933  CAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATGATGACACATTTTGGAAATGCAG 1992
Db      1749  CAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATGATGACACATTTTGGAAATGCAG 1808
Oy      1993  TCCGAAGTCCGAGAGAGCGGTGTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAAA 2052
Db      1809  TCCGAAGTCCGAGAGAGCGGTGTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAAA 1868
Oy      2053  AGAAGAGATCAGAGTGAATCATATCATATCATCAGACAGATCAAGGCTTGTGCTACGTAG 2112
Db      1869  AGAAGAGATCAGAGTGAATCATATCATATCATCAGACAGATCAAGGCTTGTGCTACGTAG 1928
Oy      2113  TCTACACAGAGAGATTCAGGCAATTTACCTTGCAGTGGGTGGAACATGGGTTCAATCA 2172
Db      1929  TCTACACAGAGAGATTCAGGCAATTTACCTTGCAGTGGGTGGAACATGGGTTCAATCA 1988
Oy      2173  AACTCTTTTAAAGTAAACCTCGAGATCATGACACAGAGCTTTGGAAAGAACTTTCTCA 2232
Db      1989  AACTCTTTTAAAGTAAACCTCGAGATCATGACACAGAGCTTTGGAAAGAACTTTCTCA 2048
Oy      2233  TAAAGATGATGATGAGATGCTCTAAGACCAAGAAATGTCATATGATGATGACACTAG 2292
Db      2049  TAAAGATGATGATGAGATGCTCTAAGACCAAGAAATGTCATATGATGATGACACTAG 2108
Oy      2293  CCAGAGGCTGTGTAACAGAGCTTATGACGCTCATCAACCAACCCCAATCTCAACAGCAT 2352

```

```

Db      2109  CCAGAGGCTGTGTAACAGAGCTTATGACGCTCATCAACCAACCCCAATCTCAACAGAT 2168
Oy      2253  GGATGAGTTCGTGTAACAGATTTGAAAGAGGACCGAAACCAACGTCGGCAAGGCCAGG 2412
Db      2169  GGATGAGTTCGTGTAACAGATTTGAAAGAGGACCGAAACCAACGTCGGCAAGGCCAGG 2228
Oy      2413  ACATACCCAGAGAAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
Db      2229  ACATACCCAGAGAAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
Oy      2473  CAGAGAGCCAGCAATTTGAGAGGCAACCAAGAGTGTCTGAGCTGATTAACCTCTAGA 2532
Db      2289  CAGAGAGCCAGCAATTTGAGAGGCAACCAAGAGTGTCTGAGCTGATTAACCTCTAGA 2348
Oy      2533  AACCTCAACAGAGTAAAGTCTGCTGACATTAACCTGAAACCAATGCAATATATCAT 2592
Db      2349  AACCTCAACAGAGTAAAGTCTGCTGACATTAACCTGAAACCAATGCAATATATCAT 2408
Oy      2593  GAACTTTTTCATGAGCATTTATGATGATGTTTACAAATGATGGAAATTCAGCTGAGTTCCA 2652
Db      2409  GAACTTTTTCATGAGCATTTATGATGATGTTTACAAATGATGGAAATTCAGCTGAGTTCCA 2468
Oy      2653  CCATTTATTAATTAATTCATGAGTAACTTTCTTAATAGCTTTT 2700
Db      2469  CCAATTTAATTAATTAATTCATGAGTAACTTTCTTAATAGCTTTT 2516

RESULT 6
PCT-US94-10151A-53
; Sequence 53, Application PC/US9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESS: FLEHR HOBBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10151A
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: PP-58750-PC/PAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FRT UR
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; PCT-US94-10151A-53

Query Match 92.6%; Score 2508; DB 5; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

193 CTGACGATGGGCTGGTTAACTAGATGTCGTCTTTCTGGGGAGTATTAATTACAC 252
9 CTGACGATGGGCTGGTTAACTAGATGTCGTCTTTCTGGGGAGTATTAATTACAC 68
253 AAGAGCAATCTATCAGATGGGAACAATGTGCCAGGCTGAATTTCTCAACAAGA 312
69 AAGAGCAATCTATCAGATGGGAACAATGTGCCAGGCTGAATTTCTCAACAAGA 128
313 AATGTTGAATCAACATATGTATCACTTCAATGGCTTGGCCAAAGCTCCAGTATCA 372
129 AATGTTGAATCAACATATGTATCACTTCAATGGCTTGGCCAAAGCTCCAGTATCA 188
373 TACCTTCTTTGGATGAGGAACGAGTGGCTGTATGTTGGACAAAGATCACATAT 432
189 TACCTTCTTTGGATGAGGAACGAGTGGCTGTATGTTGGACAAAGATCACATAT 248
433 TTCATTGCACTGGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTAAC 492
249 TTCATTGCACTGGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTAAC 308
493 CAGAGAGATGATGCAAGTGGGCTGAAAAGACATCTGAAGAATGTCTAATTTCAT 552
309 CAGAGAGATGATGCAAGTGGGCTGAAAAGACATCTGAAGAATGTCTAATTTCAT 368
553 CAGGTAATTAAGCATATATCAAGTCACTGTGTAAGCTGTGAAACGGGGCTTTTCA 612
369 CAGGTAATTAAGCATATATCAAGTCACTGTGTAAGCTGTGAAACGGGGCTTTTCA 428
613 TCCAAATTTGCACTTACATTTGAATTTGCAATCATCTGTGAGACATATTTTAACTG 672
429 TCCAAATTTGCACTTACATTTGAATTTGCAATCATCTGTGAGACATATTTTAACTG 488
673 GAACCTACATTTTGAAGGCGCGTGGGAAGTCCATATGACCCCTAACCTGTGACGC 732
489 GAACCTACATTTTGAAGGCGCGTGGGAAGTCCATATGACCCCTAACCTGTGACGC 548
733 ATCCCTTTTAAATAGATGGAATTTATCTGTGAATCTGAGCTGATTTTATGGGCGAGA 792
549 ATCCCTTTTAAATAGATGGAATTTATCTGTGAATCTGAGCTGATTTTATGGGCGAGA 608
793 CTTTGCTATCTTCCGAATCTTTGGGCGACAACCCCAATCAGACAGACAGATATTC 852
609 CTTTGCTATCTTCCGAATCTTTGGGCGACAACCCCAATCAGACAGACAGATATTC 668
853 CAGGTGGCTCAATGATCCAAAGTTCATATGTGGCCCACTCATCTCAGAGATGACAATCC 912
669 CAGGTGGCTCAATGATCCAAAGTTCATATGTGGCCCACTCATCTCAGAGATGACAATCC 728
913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACACTCTGG 972
729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACACTCTGG 788
973 AAAAATCTACACGCTAGATAGGTCAATATGCAAGAAATGACTTTTGAAGGCGACAAG 1032
789 AAAAATCTACACGCTAGATAGGTCAATATGCAAGAAATGACTTTTGAAGGCGACAAG 848
1033 TCTGTGAAATTAATGGAACAATTCCTCAAGCTGCTGATTTGGTGTGCGACAGTCC 1092
849 TCTGTGAAATTAATGGAACAATTCCTCAAGCTGCTGATTTGGTGTGCGACAGTCC 908
1093 AAATGGCATGACACTATTTTGTATGAACTGCAAGATGTATTTCTAATGAACTTTAAAGA 1152
909 AAATGGCATGACACTATTTTGTATGAACTGCAAGATGTATTTCTAATGAACTTTAAAGA 968
1153 TCTTAAAAATCCAGTTGTATATGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC 1212
969 TCTTAAAAATCCAGTTGTATATGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC 1028
1213 AGCCGTGTATATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
1029 AGCCGTGTATATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088

1273 CAGGATGAGACCCCACTATCAATGGGTCTTATCAAGAAAGATGCCCTATCCAGGCC 1332
1089 CAGGATGAGACCCCACTATCAATGGGTCTTATCAAGAAAGATGCCCTATCCAGGCC 1148
1333 AGGAATCTGTCCAGCAAAAATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1392
1149 AGGAATCTGTCCAGCAAAAATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1208
1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTAATCAATCAATTTGCTATGAACA 1452
1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTAATCAATCAATTTGCTATGAACA 1268
1453 TCGCCCAATATGATCAAAAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
1269 TCGCCCAATATGATCAAAAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1328
1513 AGTGAATCAGAAATGGAACAGTATGATGATGATGATGATGATGATGATGATGATGATG 1572
1329 AGTGAATCAGAAATGGAACAGTATGATGATGATGATGATGATGATGATGATGATGATG 1388
1573 TCTTAAAGTATTCATTTCCAAAGAGACTTGGATGATTTAAGAAAGGTTCTGTGGA 1632
1389 TCTTAAAGTATTCATTTCCAAAGAGACTTGGATGATTTAAGAAAGGTTCTGTGGA 1448
1633 AGAAATGACAGTATTTTGGGAACCGACTGTATTTAGCAATGAGCTTTCATTAAGCA 1692
1449 AGAAATGACAGTATTTTGGGAACCGACTGTATTTAGCAATGAGCTTTCATTAAGCA 1508
1693 GCAACATATATATTTGTTTCAAGGCTGGGGTTGCCAGTCTCCTTTACACCGGTGGA 1752
1509 GCAACATATATATTTGTTTCAAGGCTGGGGTTGCCAGTCTCCTTTACACCGGTGGA 1568
1753 TATTTACGGGAAGAGGTGTGCTGATGTTGTGCTGCGCCGAGACCTTACCTGTGCGGA 1812
1569 TATTTACGGGAAGAGGTGTGCTGATGTTGTGCTGCGCCGAGACCTTACCTGTGCGGA 1628
1813 TGGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAAGATAT 1872
1629 TGGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAAGATAT 1688
1873 AAGAAATGGAACCCCACTGATCTACTGTTCAGACTTACCATGATATATCAATGGCGCA 1932
1689 AAGAAATGGAACCCCACTGATCTACTGTTCAGACTTACCATGATATATCAATGGCGCA 1748
1933 CAGCCCTGAAGAGAAATCATCTATGTGTGAGAAATGTATGACATTTTGGATGAG 1992
1749 CAGCCCTGAAGAGAAATCATCTATGTGTGAGAAATGTATGACATTTTGGATGAG 1808
1993 TCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAATGAAGCGGAA 2052
1809 TCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAATGAAGCGGAA 1868
2053 AAGAAATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2112
1869 AAGAAATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1928
2113 TCTACACAGAAAGATTCAGGCAATTAATCTGTGCATGCGGTGGAACATGGGTTCAATCA 2172
1929 TCTACACAGAAAGATTCAGGCAATTAATCTGTGCATGCGGTGGAACATGGGTTCAATCA 1988
2173 AACTCTTCTTAAAGTAACTCTGGAAGTCAATTTGACACAGAGATTTTGAAGAACTTTTCA 2232
1989 AACTCTTCTTAAAGTAACTCTGGAAGTCAATTTGACACAGAGATTTTGAAGAACTTTTCA 2048
2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2292
2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2108
2293 CCAGAAAGTCTGTATCAGAGACTTCATGACAGTCAATCAACCCCAATCTCAACAGAT 2352
2109 CCAGAAAGTCTGTATCAGAGACTTCATGACAGTCAATCAACCCCAATCTCAACAGAT 2168
2353 GATGATGTTCTGTGAACAAGTTTGAAGAAAGGAGCGAACAACAGTGGCAAGGCGAGG 2412

Db 2169 GATAGATTCTGTGAACAAGTTTGGAAAAAGGACCGAAACAACTCCGCAAGCCGAG 2228
Qy 2413 ACATACCCAGGAGAACAGTAACAATGGAAGCACTTACAGAAAATAGAAAGGTAGAA 2472
Db 2229 ACATACCCAGGAGAACAGTAACAATGGAAGCACTTACAGAAAATAGAAAGGTAGAA 2288
Qy 2473 CAGAGAGCCAGCAATTTGAGAGGCACCAGAGTGTGAGCTGATTAACCTTGA 2532
Db 2289 CAGAGAGCCAGCAATTTGAGAGGCACCAGAGTGTGAGCTGATTAACCTTGA 2348
Qy 2533 AACCTCAACAGTAAGAACTTGGCTTACAGCAATTACTGAAAAACAAATGCAATTAAT 2552
Db 2349 AACCTCAACAGTAAGAACTTGGCTTACAGCAATTACTGAAAAACAAATGCAATTAAT 2408
Qy 2593 GAACCTTTTTCATGAGCATTAATGTGATGTTTACATGAGTGGAAAAATTGAGCTGAGTCA 2652
Db 2409 GAACCTTTTTCATGAGCATTAATGTGATGTTTACATGAGTGGAAAAATTGAGCTGAGTCA 2468
Qy 2653 CCAATTATTAATTAATTCATGAGTAATCTTCTTAATAGCTTTT 2700
Db 2469 CCAATTATTAATTAATTCATGAGTAATCTTCTTAATAGCTTTT 2516

RESULT 7

US-08-136-922-1
; Sequence 1, Application US/08136922
; Patent No. 5416197

GENERAL INFORMATION:

APPLICANT: Raper, Jonathan A.
APPLICANT: Luo, Yuling
TITLE OF INVENTION: Compositions Which Regulate Neural
TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5416197x18
STREET: One Liberty Place
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,922

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-1428

TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both

MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 50..1480

US-08-136-922-1

Query Match 46.3%; Score 1253; DB 1; Length 1481;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1303; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 904 TGACAAATCTGAAAGATGACAAAGTATCTTTTCTTCGGTAAATGCAATAGATGAGA 963
Db 58 TGAACAATCTGAAAGATGACAAAGTATCTTTTCTTCGGTAAATGCAATAGATGAGA 117
Qy 964 ACATCTTGAAAAAGCTCTCACTGATGAATAGTCAATATGCAAGATGACTTTGAGG 1023
Db 118 ACATCTTGAAAAAGCTCTCACTGATGAATAGTCAATATGCAAGATGACTTTGAGG 1177
Qy 1024 GCAAGAAAGTGTGTAATTAATGGAACAATCTCAAGCTGCTGATTTGCTCAGT 1083
Db 178 GCACAGAAAGTGTGTAATTAATGGAACAATCTCAAGCTGCTGATTTGCTCAGT 237
Qy 1084 GCCAGTTCAAAATGCGATTGACACTCATTTTGTGATGAATGAGATGATATCTTAATGAA 1143
Db 238 GCACAGTTCAAAATGCGATTGACACTCATTTTGTGATGAATGAGATGATATCTTAATGAA 297
Qy 1144 CTTTAAAGATCTTAAATATCCAGTTGTATATGAGTGTTCAGACTTCAGTAAATTTT 1203
Db 298 CTTTAAAGATCTTAAATATCCAGTTGTATATGAGTGTTCAGACTTCAGTAAATTTT 357
Qy 1204 CAAGGATCAGCGGTGTATATGATAGATGATGAGTGAAGAGGTGTTCTTGATTC 1263
Db 358 CAAGGATCAGCGGTGTATATGATAGATGATGAGTGAAGAGGTGTTCTTGATTC 417
Qy 1264 ATATGCCACAGAGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTTA 1323
Db 418 ATATGCCACAGAGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGTCCCTTA 477
Qy 1324 TCACGGCCAGAACTTGTCCAGCAAAACATTTGTGTTTGACTCTCAAAAGACT 1383
Db 478 TCACGGCCAGAACTTGTCCAGCAAAACATTTGTGTTTGACTCTCAAAAGACT 537
Qy 1384 TCCGATGATGTTATTAATCTTTGCAAGATGATCCAGCATGATGATCAATGATTC 1443
Db 538 TCCGATGATGTTATTAATCTTTGCAAGATGATCCAGCATGATGATCAATGATTC 597
Qy 1444 TATGAACAATCGCCCAATAGTATCAAAACGATGATTAATTAATTAATTAATTAAT 1503
Db 598 TATGAACAATCGCCCAATAGTATCAAAACGATGATTAATTAATTAATTAATTAAT 657
Qy 1504 GTAGACCGAGTGTATGACAGAAAGTGAACGATGATGATGATGATGATGATGAT 1563
Db 658 GTAGACCGAGTGTATGACAGAAAGTGAACGATGATGATGATGATGATGATGAT 717
Qy 1564 TGGGACCGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1623
Db 718 TGGGACCGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 777
Qy 1624 TCGCTGGAAGAAATGACAGTTTTCGGAACGAGTGTATTTTCAAGCAATGAGCTTTC 1683
Db 778 TCGCTGGAAGAAATGACAGTTTTCGGAACGAGTGTATTTTCAAGCAATGAGCTTTC 837
Qy 1684 CACTAAGCAGCAACAATATATATTTGTTCAACGCTGAGGTGCCAGCTCCCTTTACA 1743
Db 838 CACTAAGCAGCAACAATATATATTTGTTCAACGCTGAGGTGCCAGCTCCCTTTACA 897
Qy 1744 CCGGTGTATATTTAATGAGAAAGCTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTG 1803
Db 898 CCGGTGTATATTTAATGAGAAAGCTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTG 957
Qy 1804 TGGTGGGATGATTTCTGCAATGTTCTGCTATTTTCCCACTGCAAAAGAGGCAAAAGC 1863
Db 958 TGGTGGGATGATTTCTGCAATGTTCTGCTATTTTCCCACTGCAAAAGAGGCAAAAGC 1017
Qy 1864 ACAAGATATTAAGAAATGAGAACCACTGACTGCTGTTCAAGCTTACACCATGATATCA 1923
Db 1018 ACAAGATATTAAGAAATGAGAACCACTGACTGCTGTTCAAGCTTACACCATGATATCA 1077
Qy 1924 CCATGGCCACAGCCCTGAAAGAGAAATCATCTATGCTGTAGAAATATGACATTTT 1983
Db 1078 CCATGGCCACAGCCCTGAAAGAGAAATCATCTATGCTGTAGAAATATGACATTTT 1137

QY 1984 GGAATGCACTCCGAAGTCCAGAGCGCTGTCTATTGGCAATTCCAGAGGCGAATGA 2043
DB 1138 GGAATGCACTCCGAAGTCCAGAGCGCTGTCTATTGGCAATTCCAGAGGCGAATGA 1197
QY 2044 AGAGCGAAGAGAGATCAGATGATGATCTATCATCAGACAGATCAAGGCTTCT 2103
DB 1198 AGAGCGAAGAGAGATCAGATGATGATCTATCATCAGACAGATCAAGGCTTCT 1257
QY 2104 GGTACGATCTTACACAGAGAGATCAGACATTAAGCTGCGATGCGGAGACATG 2163
DB 1258 GGTACGATCTTACACAGAGAGATCAGACATTAAGCTGCGATGCGGAGACATG 1317
QY 2164 GTTCATACAACTCTTCTTAAGTAACTCTGGAAGTCAATGACA 2207
DB 1318 GTTCATACAACTCTTCTTAAGTAACTCTGGAAGTCAATGACA 1361

RESULT 8
US-08-786-531B-5/c
Sequence 5, Application US/08786531B
Patent No. 6541197
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suning
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7160 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-5

Query Match 1.1%; Score 31; DB 4; Length 7160;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTATTTATCGATGTTTAAACAAGCTT 31
DB 4179 AATCTTTATTTATCGATGTTTAAACAAGCTT 4149

RESULT 9
US-08-786-531B-6/c
Sequence 6, Application US/08786531B
Patent No. 6541197
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suning
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 7235 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-6

Query Match 1.1%; Score 31; DB 4; Length 7235;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTATTTATCGATGTTTAAACAAGCTT 31
DB 4254 AATCTTTATTTATCGATGTTTAAACAAGCTT 4224

ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgttri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 7352 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-4

Query Match 0.8%; Score 22; DB 4; Length 7352;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
DB 4371 AATCTTTATTTATCGATGTT 4350

RESULT 11
US-08-786-531B-1/C
Sequence 1, Application US/08786531B
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgttri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-1

Query Match 0.8%; Score 22; DB 4; Length 7353;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
DB 4372 AATCTTTATTTATCGATGTT 4351

RESULT 12
US-09-391-741A-23
Sequence 23, Application US/09391741A
Patent No. 655732
GENERAL INFORMATION:
APPLICANT: Duvick, Jonathan P.
APPLICANT: Sharma, Yogesh Kumar
TITLE OF INVENTION: Rac-Like Genes and Methods of Use
FILE REFERENCE: 0866D
CURRENT APPLICATION NUMBER: US/09/391,741A
CURRENT FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: 60/111,919
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/100,284
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 09/391,741
PRIOR FILING DATE: 1999-09-08
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 1062
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (172)...(807)
US-09-391-741A-23

Query Match 0.8%; Score 21; DB 4; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCAGCGCTCCGGAG 77
DB 3 GTCGACCCAGCGCTCCGGAG 23

RESULT 13
US-09-391-741A-33
Sequence 33, Application US/09391741A
Patent No. 655732
GENERAL INFORMATION:
APPLICANT: Duvick, Jonathan P.

APPLICANT: Sharma, Yogesh Kumar
TITLE OF INVENTION: Rac-Like Genes and Methods of Use
FILE REFERENCE: 0866D
CURRENT APPLICATION NUMBER: US/09/391,741A
CURRENT FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: 60/111,919
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/100,284
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 09/391,741
PRIOR FILING DATE: 1999-09-08
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 1062
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (172)..(807)
US-09-391-741A-33

Query Match 0.8%; Score 21; DB 4; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGAG 77
|||||
DB 3 GTCGACCCACGCGTCCGGAG 23

RESULT 14

US-09-620-312D-33
Sequence 33, Application US/09620312D

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIPB
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 33
LENGTH: 1181
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)..(874)
US-09-620-312D-33

Query Match 0.8%; Score 21; DB 4; Length 1181;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGAG 77
|||||
DB 59 GTCGACCCACGCGTCCGGAG 79

RESULT 15

US-08-984-288-1
Sequence 1, Application US/08984288

GENERAL INFORMATION:
APPLICANT: BERGSM, DEBK
APPLICANT: ELLIS, CATHERINE
TITLE OF INVENTION: NOVEL HUMAN 11CB SPLICE V
TITLE OF INVENTION: ARIANT
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,288
FILING DATE: 03-DEC-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,763
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: P50599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-984-288-1

Query Match 0.8%; Score 21; DB 3; Length 1385;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGAG 77
|||||
DB 50 GTCGACCCACGCGTCCGGAG 70

RESULT 16

US-09-620-312D-664
Sequence 664, Application US/09620312D

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John, Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 664
LENGTH: 1572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (330)..(704)
US-09-620-312D-664

Query Match 0.8%; Score 21; DB 4; Length 1572;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACCGCGTCGGGAG 77
Db 280 GTGACCCACCGCGTCGGGAG 300

RESULT 17
US-08-815-176-2
Sequence 2, Application US/08815176
Patent No. 5874224
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Diegidio, Tony
TITLE OF INVENTION: NOVEL EGF RECEPTOR BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/815,176
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0236 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: PENNOT03
CLONE: 1291904
US-08-815-176-2

Query Match 0.8%; Score 21; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACCGCGTCGGGAG 77
Db 132 GTGACCCACCGCGTCGGGAG 152

RESULT 18
US-09-197-344-2
Sequence 2, Application US/09197344
Patent No. 6417329
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Diegidio, Tony
TITLE OF INVENTION: NOVEL EGF RECEPTOR BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,344
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/815,176
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0236 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: PENNOT03
CLONE: 1291904
US-09-197-344-2

Query Match 0.8%; Score 21; DB 4; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 132 GTGACCCACGCGTCCGGAG 152

RESULT 19

US-09-324-455-1
Sequence 1, Application US/09324455
Patent No. 6326481
GENERAL INFORMATION:
APPLICANT: Yowe, David
TITLE OF INVENTION: NOVEL MOLECULES OF THE AIP-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09324,455
FILING DATE: 02-JUN-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/087,761
FILING DATE: 02-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Weiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/069001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2026 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-324-455-1

Query Match 0.8%; Score 21; DB 4; Length 2026;
Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 11 GTGACCCACGCGTCCGGAG 31

RESULT 20
US-08-850-961-31/c
Sequence 31, Application US/08850961
Patent No. 6013517
GENERAL INFORMATION:
APPLICANT: Respass, James G.
APPLICANT: De Polo, Nicholas J.
APPLICANT: Canada, Sunil
APPLICANT: Sauter, Sybille
APPLICANT: Bodner, Mordechai
APPLICANT: Driver, David A.
TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation, Intellectual Property - R440
STREET: P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94662-8097

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,961
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kruse, No. 6013517man J.
REGISTRATION NUMBER: 35,235
REFERENCE/DOCKET NUMBER: 930049.424C4 / 1147.005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-3520
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-850-961-31

Query Match 0.7%; Score 20; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
Db 20 AATCTTTATTTATCGATG 1

RESULT 21
US-09-132-541-5/c
Sequence 5, Application US/09132541A
Patent No. 6114113
GENERAL INFORMATION:
APPLICANT: MCLAUGHLIN-TAYLOR, Elizabeth
APPLICANT: KRUGER, Mark
APPLICANT: LUNDAK, Cheryl
APPLICANT: KILLION, Catherine
TITLE OF INVENTION: HIGH EFFICIENCY GENETIC MODIFICATION METHODS
FILE REFERENCE: 1386.002
CURRENT APPLICATION NUMBER: US/09/132,541A
CURRENT FILING DATE: 1998-08-11
EARLIER APPLICATION NUMBER: 60/055,453
EARLIER FILING DATE: 1997-08-11
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 30
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: This information is not available
US-09-132-541-5

Query Match 0.7%; Score 20; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
Db 20 AATCTTTATTTATCGATG 1

RESULT 22
US-09-479-776-31/c
Sequence 31, Application US/09479776
Patent No. 6333195
GENERAL INFORMATION:
APPLICANT: Respess, James G.
De Polo, Nicholas J.
Chada, Sunil
Sauter, Sybille
Bodner, Mordechai
Driver, David A.
TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: INTELLECTUAL PROPERTY-R440
P.O. BOX 8097
CITY: EMERYVILLE
STATE: CA
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,776
FILING DATE: 07-Jan-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: KRUSE, NORMAN J.
REGISTRATION NUMBER: 35,235
REFERENCE/DOCKET NUMBER: 930049.424C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)622-4900
FAX: (206)682-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-479-776-31

Query Match 0.7%; Score 20; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATTCGATG 20
Db 20 AATCTTTATTTATTCGATG 1

RESULT 23
US-09-244-111-5
Sequence 5, Application US/09244111
Patent No. 6566498
GENERAL INFORMATION:
APPLICANT: Ni, et al.
TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
FILE REFERENCE: PF391
CURRENT APPLICATION NUMBER: US/09/244,111
CURRENT FILING DATE: 1999-02-04
EARLIER APPLICATION NUMBER: 60/073,961
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 840
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: CDS
LOCATION: (115)..(603)
US-09-244-111-5

Query Match 0.7%; Score 20; DB 4; Length 840;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGGA 76
Db 4 GTGACCCACGCGTCGGGGA 23

RESULT 24
US-09-391-741A-9
Sequence 9, Application US/09391741A
Patent No. 655732
GENERAL INFORMATION:
APPLICANT: Davick, Jonathan P.
Sharma, Yogesh Kumar
TITLE OF INVENTION: Rac-Like Genes and Methods of Use
FILE REFERENCE: 0866D
CURRENT APPLICATION NUMBER: US/09/391,741A
CURRENT FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: 60/111,919
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/100,284
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 09/391,741
PRIOR FILING DATE: 1999-09-08
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 1059
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (169)..(804)
US-09-391-741A-9

Query Match 0.7%; Score 20; DB 4; Length 1059;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGCGTCGGGAG 77
Db 1 TCGACCCACGCGTCGGGAG 20

RESULT 25
US-09-620-312D-562
Sequence 562, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
Lin, Chenghua
APPLICANT: Asundi, Vinod
Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungang
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_fl_genes Version 1.0
SEQ ID NO: 562
LENGTH: 1168
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (148)..(615)
US-09-620-312D-562

Query Match 0.7%; Score 20; DB 4; Length 1168;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGA 76
DB 25 GTGACCCACGCGTCCGGA 44

RESULT 26
US-08-985-950-5
Sequence 5, Application US/08985950
Patent No. 6140076

GENERAL INFORMATION:

APPLICANT: Adema, Gosee Jan
TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/032,252
FILING DATE: 06-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0670K
TELECOMMUNICATION INFORMATION:

TELEPHONE: (650)852-9196
TELEFAX: (650)496-1204
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1279 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..1015
NAME/KEY: misc_feature
LOCATION: 1247
OTHER INFORMATION: /note= "nucleotide 1247 designated
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 218..1015
US-08-985-950-5

Query Match 0.7%; Score 20; DB 3; Length 1279;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGA 76
DB 20 GTGACCCACGCGTCCGGA 39

RESULT 27
US-09-546-049-5
Sequence 5, Application US/09546049
Patent No. 6479638

GENERAL INFORMATION:

APPLICANT: Adema, Gosee Jan
Meygaard, Lunde

Gorman, Daniel M.
McClanahan, Terrill K.

Zurawski, Sandra M.
Zurawski, Gerard

Lanier, Lewis L.
Phillips Jr., Joseph H.

TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
Related Reagents

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/546,049
FILING DATE: 10-APR-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0670K
TELECOMMUNICATION INFORMATION:

TELEPHONE: (650)852-9196
TELEFAX: (650)496-1204
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 1279 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..1015
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1247
OTHER INFORMATION: /note= "nucleotide 1247 designated
C, but may be C or T"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 218..1015
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-546-049-5

Query Match 0.7%; Score 20; DB 4; Length 1279;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACCGCTCCGGGA 76
|||||
Db 20 GTCGACCCACCGCTCCGGGA 39

RESULT 28
US-09-336-536-1
Sequence 1, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
US-09-336-536-1

Query Match 0.7%; Score 20; DB 4; Length 1338;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACCGCTCCGGGA 76
|||||
Db 1 GTCGACCCACCGCTCCGGGA 20

RESULT 29
US-08-999-774A-11
Sequence 11, Application US/08999774A
Patent No. 6274312
GENERAL INFORMATION:
APPLICANT: Gish, Kurt C.
APPLICANT: Seghezzi, Wolfgang
APPLICANT: Shanahan, Frances
APPLICANT: Lees, Emma M.
APPLICANT: McClanahan, Terrill K.
TITLE OF INVENTION: Intracellular Regulatory Molecules;
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,774A
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,818
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0646
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-9196
TELEFAX: (650) 496-1200
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1503 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1503
US-08-999-774A-11

Query Match 0.7%; Score 20; DB 3; Length 1503;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACCGCTCCGGGA 76
|||||
Db 1 GTCGACCCACCGCTCCGGGA 20

RESULT 30
US-09-482-273-96
Sequence 96, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 96
LENGTH: 1772
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-96

Query Match 0.7%; Score 20; DB 4; Length 1772;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 31

US-09-620-312D-7
Sequence 7, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pc_FL_genes Version 1.0

SEQ ID NO 7

LENGTH: 1920

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (120)..(1394)

US-09-620-312D-7

Query Match

Best Local Similarity 0.7%; Score 20; DB 4; Length 1920;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGGA 76

|||||

DB 21 GTCGACCCACGCGTCCGGGA 40

RESULT 32

US-09-489-847-113

Sequence 113, Application US/09489847

Patent No. 6476195

GENERAL INFORMATION:

APPLICANT: Rosen et al

TITLE OF INVENTION: 98 Human Secreted Proteins

FILE REFERENCE: P2031p1

CURRENT APPLICATION NUMBER: US/09/489,847

CURRENT FILING DATE: 2000-01-24

EARLIER APPLICATION NUMBER: PCT/US99/17130

EARLIER FILING DATE: 1999-07-29

EARLIER APPLICATION NUMBER: 60/094,657

EARLIER FILING DATE: 1998-07-30

EARLIER APPLICATION NUMBER: 60/095,486

EARLIER FILING DATE: 1998-08-05

EARLIER APPLICATION NUMBER: 60/096,319

EARLIER FILING DATE: 1998-08-12

EARLIER APPLICATION NUMBER: 60/095,454

EARLIER FILING DATE: 1998-08-06

EARLIER APPLICATION NUMBER: 60/095,455

EARLIER FILING DATE: 1998-08-06

NUMBER OF SEQ ID NOS: 376

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 113

LENGTH: 2214

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (1289)

OTHER INFORMATION: n equals a,t,c,g, or c

US-09-489-847-113

Query Match

Best Local Similarity 0.7%; Score 20; DB 4; Length 2214;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGGA 76

|||||

DB 12 GTCGACCCACGCGTCCGGGA 31

RESULT 33

US-09-620-312D-1046

Sequence 1046, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pc_FL_genes Version 1.0

SEQ ID NO 1046

LENGTH: 2225

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (260)..(1906)

US-09-620-312D-1046

Query Match

Best Local Similarity 0.7%; Score 20; DB 4; Length 2225;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 35 GTGACCCACGGCTCCGGGA 54

RESULT 34

US-09-489-847-30
; Sequence 30, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-30

Query Match Best Local Similarity 0.7%; Score 20; DB 4; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 12 GTGACCCACGGCTCCGGGA 31

RESULT 35

US-09-221-235-1
; Sequence 1, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-235-1

Query Match Best Local Similarity 0.7%; Score 20; DB 3; Length 4137;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77
|||||
Db 1 TCGACCCACGGCTCCGGAG 20

RESULT 36

US-09-221-928-1
; Sequence 1, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-928-1

Query Match Best Local Similarity 0.7%; Score 20; DB 3; Length 4137;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77
|||||
Db 1 TCGACCCACGGCTCCGGAG 20

RESULT 37

US-09-221-527-1
; Sequence 1, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-527-1

Query Match Best Local Similarity 0.7%; Score 20; DB 3; Length 4137;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77
|||||

DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 38

US-09-221-236-1
; Sequence 1, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-236-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 39

US-09-221-416-1
; Sequence 1, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-416-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 40

US-09-221-245-1
; Sequence 1, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-245-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 41

US-09-163-115-1
; Sequence 1, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-163-115-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 42

US-09-221-528-1
; Sequence 1, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050

```
;; CURRENT APPLICATION NUMBER: US/09/221,528
;; CURRENT FILING DATE: 1998-12-28
;; EARLIER APPLICATION NUMBER: 09/163,115
;; EARLIER FILING DATE: 1998-09-29
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 4137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
;; NAME/KEY: CDS
;; LOCATION: (297)..(1202)
US-09-221-528-1

Query Match      0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGGCGTCGGGAG 77
DB 1 TCGACCCACGGCGTCGGGAG 20

RESULT 43
US-09-593-553-1
;; Sequence 1, Application US/09593553
;; Patent No. 6200770
;; GENERAL INFORMATION:
;; APPLICANT: Acton, Susan
;; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
;; FILE REFERENCE: NMI-050
;; CURRENT APPLICATION NUMBER: US/09/593,553
;; CURRENT FILING DATE: 2000-06-14
;; PRIOR APPLICATION NUMBER: 09/163,115
;; PRIOR FILING DATE: 1998-09-28
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 4137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
;; NAME/KEY: CDS
;; LOCATION: (297)..(1202)
US-09-593-553-1

Query Match      0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGGCGTCGGGAG 77
DB 1 TCGACCCACGGCGTCGGGAG 20

RESULT 44
US-09-221-237-1
;; Sequence 1, Application US/09221237
;; Patent No. 6214597
;; GENERAL INFORMATION:
;; APPLICANT: Acton, Susan
;; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
;; FILE REFERENCE: MNI-050
;; CURRENT APPLICATION NUMBER: US/09/221,237
;; CURRENT FILING DATE: 1998-12-28
;; EARLIER APPLICATION NUMBER: 09/163,115
;; EARLIER FILING DATE: 1998-09-29
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: Patentln Ver. 2.0
```

```
;; SEQ ID NO 1
;; LENGTH: 4137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
;; NAME/KEY: CDS
;; LOCATION: (297)..(1202)
US-09-221-237-1

Query Match      0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGGCGTCGGGAG 77
DB 1 TCGACCCACGGCGTCGGGAG 20

RESULT 45
US-08-654-737B-3/C
;; Sequence 3, Application US/08654737B
;; Patent No. 6274136
;; GENERAL INFORMATION:
;; APPLICANT: University of Southern California
;; APPLICANT: Weiner, Leslie P.
;; APPLICANT: McMillan, Mtnie
;; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
;; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
;; FILE REFERENCE: 13761-703-00 US
;; CURRENT APPLICATION NUMBER: US/08/654,737B
;; CURRENT FILING DATE: 1996-05-29
;; NUMBER OF SEQ ID NOS: 12
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 5865
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Retroviral vector derived from Moloney Murine
;; OTHER INFORMATION: Leukemia Virus
US-08-654-737B-3

Query Match      0.7%; Score 20; DB 3; Length 5865;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
DB 2740 AATCTTTATTTATCGATG 2721

RESULT 46
US-09-654-449-1/c
;; Sequence 1, Application US/09654449
;; Patent No. 6564061
;; GENERAL INFORMATION:
;; APPLICANT: Palmer, Michelle A.J.
;; APPLICANT: Gee, Melissa
;; APPLICANT: Tilletson, Bonnie
;; APPLICANT: Chang, Xiao-jia
;; TITLE OF INVENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orphan
;; TITLE OF INVENTION: Receptors by Reporter Enzyme Mutant Complementation
;; FILE REFERENCE: 4085-226-27
;; CURRENT APPLICATION NUMBER: US/09/654,449
;; CURRENT FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: US 60/180,669
;; PRIOR FILING DATE: 2000-02-07
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 6700
```

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Nucleotide sequence for pICAST ALC
NAME/KEY: CDS
LOCATION: (1457)...(4486)
US-09-654-449-1

Query Match 0.7%; Score 20; DB 4; Length 6700;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
DB 5876 AATCTTTATTTATCGATG 5857

RESULT 47
US-09-654-449-3/c
Sequence 3, Application US/09654449
Patent No. 6564061
GENERAL INFORMATION:
APPLICANT: Palmer, Michelle A.J.
APPLICANT: Gee, Melissa
APPLICANT: Tillotson, Bonnie
APPLICANT: Chang, Xiao-Jia
TITLE OF INVENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orphan
FILE REFERENCE: 4085-226-27
CURRENT APPLICATION NUMBER: US/09/654,449
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: US 60/180,669
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 8518
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Nucleotide sequence for pICAST ALN
US-09-654-449-3

Query Match 0.7%; Score 20; DB 4; Length 8518;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
DB 5877 AATCTTTATTTATCGATG 5858

RESULT 48
US-08-487-283A-4/c
Sequence 4, Application US/08487283A
Patent No. 6355245
GENERAL INFORMATION:
APPLICANT: Evans, Mark J.
APPLICANT: Mattis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park (Alexion)

CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8540 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Apex-3p Eukaryotic
DESCRIPTION: Expression Vector
US-08-487-283A-4

Query Match 0.7%; Score 20; DB 4; Length 8540;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
DB 3509 AATCTTTATTTATCGATG 3490

RESULT 49
PCT-US96-05611A-12/c
Sequence 12, Application PC/TUS9605611A
GENERAL INFORMATION:
APPLICANT: Mueller, John P.
APPLICANT: Leonardo, Michael J.
APPLICANT: McFarland, Henry F.
APPLICANT: Mattis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Palfrey, Clara M.
APPLICANT: Squinto, Stephen P.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: Modified Myelin Protein Molecules
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05611A
FILING DATE: 02-MAY-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,644
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/431,648
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/482,114
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 8540 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Apex-3p Eukaryotic
Expression Vector
PCT-US96-05611A-12

Query Match 0.7%; Score 20; DB 5; Length 8540;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTTATCGATG 20
|||||
DB 3509 AATCTTTATTTTATCGATG 3490

RESULT 50
US-09-313-294A-2526
Sequence 2526, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalugudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2526
LENGTH: 256
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700552828H1
US-09-313-294A-2526

Query Match 0.7%; Score 19; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1587 CAATCCTAAGAGACTTG 1605
|||||
DB 58 CAATCCTAAGAGACTTG 76

RESULT 51
US-09-461-697-136
Sequence 136, Application US/09461697
Patent No. 6272974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.

APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kasuri
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 136
LENGTH: 542
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-697-136

Query Match 0.7%; Score 19; DB 3; Length 542;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCGGGA 76
|||||
DB 1 TCGACCCACGCGTCGGGA 19

RESULT 52
US-09-620-312D-920
Sequence 920, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 920
LENGTH: 571
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (63)..(425)
US-09-620-312D-920

Query Match 0.7%; Score 19; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGG 75
Db 31 GTGACCCACGCGTCCGG 49

RESULT 53
US-09-512-363-9
Sequence 9, Application US/09512363
Patent No. 6503184

GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
TITLE OF INVENTION: Tr11, Tr11SV1, and Tr11SV2
FILE REFERENCE: PF396
CURRENT APPLICATION NUMBER: US/09/512,363
CURRENT FILING DATE: 2000-02-23
EARLIER APPLICATION NUMBER: 60/063,212
EARLIER FILING DATE: 1997-10-21
EARLIER APPLICATION NUMBER: 09/176,200
EARLIER FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: 60/121,648
EARLIER FILING DATE: 1999-02-24
EARLIER APPLICATION NUMBER: 60/134,172
EARLIER FILING DATE: 1999-05-13
EARLIER APPLICATION NUMBER: 60/144,076
EARLIER FILING DATE: 1999-07-16
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 581

TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc_feature
LOCATION: (291)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (293)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (295)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (319)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (402)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (424)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (452)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (440)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (471)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (483)

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (495)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (516)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (545)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (559)
OTHER INFORMATION: n equals a, t, g or c

US-09-512-363-9

Query Match 0.7%; Score 19; DB 4; Length 581;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGG 75
Db 3 GTGACCCACGCGTCCGG 21

RESULT 54
US-09-176-200-9
Sequence 9, Application US/09176200
Patent No. 6509173

GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
TITLE OF INVENTION: Tr11, Tr11SV1, and Tr11SV2
FILE REFERENCE: PF396
CURRENT APPLICATION NUMBER: US/09/176,200
CURRENT FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: 60/063,212
EARLIER FILING DATE: 1997-10-21
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 581

TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc_feature
LOCATION: (291)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (293)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (295)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (319)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (402)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (424)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (440)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (471)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (483)

NAME/KEY: misc_feature
LOCATION: (452)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (440)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (471)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (483)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (495)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (516)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (545)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559)
OTHER INFORMATION: n equals a, t, g or c
US-09-176-200-9

Query Match 0.7%; Score 19; DB 4; Length 581;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 57 GTGACCCAGCGCTCGGG 75
Db 3 GTGACCCAGCGCTCGGG 21

RESULT 55
US-08-801-972-2
Sequence 2, Application US/08801972
Patent No. 5831018
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,972
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0216 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2264814
US-08-801-972-2

Query Match 0.7%; Score 19; DB 2; Length 648;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1615 AGAAGGTTCTGCTGGA 1633
Db 243 AGAAGGTTCTGCTGGA 261

RESULT 56
US-09-178-881-2
Sequence 2, Application US/09178881
Patent No. 6083704
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME B5
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/178,881
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/801,972
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0216 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2264814
US-09-178-881-2

Query Match 0.7%; Score 19; DB 3; Length 648;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1615 AGAAGGTTCTGCTGGA 1633
|||||

DB 243 AGAAGAGTTCTGCTGAA 261

RESULT: 57

US-09-205-258-63

Sequence 63, Application US/09205258

Patent No. 6525174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258

EARLIER FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,972

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,916

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,374

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,917

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,949

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 63
LENGTH: 780
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (738)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (776)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-63

Query Match 0.7%; Score 19; DB 4; Length 780;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCCGCCACCGCGCGG 75
DB 28 GTCCGCCACCGCGCGG 46

RESULT 58
US-09-904-615-25
Sequence 25, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 49 Human Secreted Proteins
FILE REFERENCE: P2032P1
CURRENT APPLICATION NUMBER: US/09/904,615
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/511,554
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/097,917
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 60/098,634
PRIOR FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 831
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (5)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (11)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (15)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-25

Query Match 0.7%; Score 19; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACCGCGTCCGGG 75
DB 67 GTGACCCACCGCGTCCGGG 85

RESULT 59
US-09-461-697-1
Sequence 1, Application US/09461697
Patent No. 6277974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kasumi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT FILING DATE: 1999-12-14
CURRENT APPLICATION NUMBER: US/09/461,697
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 852
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-697-1

Query Match 0.7%; Score 19; DB 3; Length 852;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 TCGACCCACCGCGTCCGGGA 76
DB 1 TCGACCCACCGCGTCCGGGA 19

RESULT 60
US-09-257-179-20
Sequence 20, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P201SP1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 932
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (43)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (458)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (465)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (494)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (501)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (515)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (582)
OTHER INFORMATION: n equals a,t,g, or c
US-09-257-179-20

Query Match 0.7%; Score 19; DB 4; Length 932;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACCGCGTCCGGG 75
DB 72 GTGACCCACCGCGTCCGGG 90

RESULT 61
US-09-512-363-3
Sequence 3, Application US/09512363
Patent No. 6503184
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
FILE REFERENCE: PR396
CURRENT APPLICATION NUMBER: US/09/512,363
CURRENT FILING DATE: 2000-02-23
EARLIER APPLICATION NUMBER: 60/063,212

EARLIER FILING DATE: 1997-10-21
EARLIER APPLICATION NUMBER: 09/176,200
EARLIER FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: 60/121,648
EARLIER FILING DATE: 1999-02-24
EARLIER APPLICATION NUMBER: 60/134,172
EARLIER FILING DATE: 1999-05-13
EARLIER APPLICATION NUMBER: 60/144,076
EARLIER FILING DATE: 1999-07-16
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 3
LENGTH: 1007
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (121)..(843)
US-09-512-363-3

Query Match 0.7%; Score 19; DB 4; Length 1007;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
Db 1 GTGACCCACGCGTCGGG 19

RESULT 62
US-09-176-200-3
Sequence 3, Application US/09176200
Patent No. 6509173
GENERAL INFORMATION:
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
FILE REFERENCE: P396
CURRENT APPLICATION NUMBER: US/09/176,200
CURRENT FILING DATE: 1998-10-21
EARLIER APPLICATION NUMBER: 60/063,212
EARLIER FILING DATE: 1997-10-21
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 3
LENGTH: 1007
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (121)..(843)
US-09-176-200-3

Query Match 0.7%; Score 19; DB 4; Length 1007;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
Db 1 GTGACCCACGCGTCGGG 19

RESULT 63
US-09-620-312D-866
Sequence 866, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aiyong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillmanast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO: 866
LENGTH: 1024
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (519)..(980)
US-09-620-312D-866

Query Match 0.7%; Score 19; DB 4; Length 1024;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
Db 14 GTGACCCACGCGTCGGG 32

RESULT 64
US-09-489-847-112
Sequence 112, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 112
LENGTH: 1037
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (936)
OTHER INFORMATION: n equals a,t,g, or c

```
FEATURE:
NAME/KEY: SITE
LOCATION: (946)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (951)
OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-112
```

```
Query Match          0.7%; Score 19; DB 4; Length 1037;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 57 GTCCAGCCACGCGTCGGG 75
Db 19 GTCCAGCCACGCGTCGGG 37
```

```
RESULT 65
US-09-372-422A-41
Sequence 41, Application US/09372422A
Patent No. 6313375
GENERAL INFORMATION:
APPLICANT: Rudolf Jung
TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
FILE REFERENCE: 0919
CURRENT APPLICATION NUMBER: US/09/372,422A
CURRENT FILING DATE: 1999-08-11
PRIOR APPLICATION NUMBER: US 60/098,692
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 41
LENGTH: 1116
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (2)...(719)
US-09-372-422A-41
```

```
Query Match          0.7%; Score 19; DB 4; Length 1116;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 57 GTCCAGCCACGCGTCGGG 75
Db 6 GTCCAGCCACGCGTCGGG 24
```

```
RESULT 66
US-09-620-312D-162
Sequence 162, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
```

```
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 162
LENGTH: 1221
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (263)..(862)
US-09-620-312D-162
```

```
Query Match          0.7%; Score 19; DB 4; Length 1221;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 57 GTCCAGCCACGCGTCGGG 75
Db 19 GTCCAGCCACGCGTCGGG 37
```

```
RESULT 67
US-09-620-312D-972
Sequence 972, Application US/09620312D
Patent No. 6569662
```

```
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 972
LENGTH: 1281
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1172)..(1074)
US-09-620-312D-972
```

```
Query Match          0.7%; Score 19; DB 4; Length 1281;
```

Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCGGG 75
|||||
DB 14 GTGACCCACGCGTCGGG 32

RESULT 68

US-09-620-312D-196

Sequence 196, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 656962e1 Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_fl_genes Version 1.0

SEQ ID NO 196

LENGTH: 1332

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (171)..(920)

US-09-620-312D-196

Query Match 0.7%; Score 19; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 24;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCGGG 75
|||||
DB 40 GTGACCCACGCGTCGGG 58

RESULT 69

US-09-620-312D-308

Sequence 308, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 656962e1 Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_fl_genes Version 1.0

SEQ ID NO 871

LENGTH: 1368

TYPE: DNA

APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_fl_genes Version 1.0
SEQ ID NO 308
LENGTH: 1354
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (226)..(1251)
US-09-620-312D-308

Query Match 0.7%; Score 19; DB 4; Length 1354;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCGGG 75
|||||
DB 129 GTGACCCACGCGTCGGG 147

RESULT 70

US-09-620-312D-871

Sequence 871, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yungqing

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 656962e1 Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT APPLICATION NUMBER: US/09/620,312D

CURRENT FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_fl_genes Version 1.0

SEQ ID NO 871

LENGTH: 1368

TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (103)..(447)
US-09-620-312D-871

Query Match 0.7%; Score 19; DB 4; Length 1368;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGG 75
Db 48 GTGACCCACGCGTCCGGG 66

RESULT 71
US-09-205-258-124
Sequence 124, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 124
LENGTH: 1390
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (498)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (499)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-124

Query Match 0.7%; Score 19; DB 4; Length 1390;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGG 75
Db 68 GTGACCCACGCGTCCGGG 86

RESULT 72
US-08-964-127-5
Sequence 5, Application US/08964127
Patent No. 6277565
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,127
FILING DATE: 06-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
US-09-964-127-5

Query Match 0.7%; Score 19; DB 3; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 1 GTGACCCACGCGTCCGGG 19

RESULT 73
US-09-496-692-5
Sequence 5, Application US/09496692
Patent No. 6313271
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/496,692
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/964,127
FILING DATE: 06-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
US-09-496-692-5

Query Match 0.7%; Score 19; DB 4; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
|||||
DB 1 GTGACCCACGCGTCCGGG 19

RESULT 74
US-10-000-273-5
Sequence 5, Application US/10000273
Patent No. 6573057
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/000,273
FILING DATE: 02-NOV-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/964,127
FILING DATE: 06-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-000-273-5

Query Match 0.7%; Score 19; DB 4; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;

1

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGG 75
 |||||
 Db 1 GTCGACCCACGCGTCCGGG 19

RESULT 75

US-09-118-442-5
 ; Sequence 5, Application US/09118442B
 ; Patent No. 6197561
 ; GENERAL INFORMATION:
 ; APPLICANT: Martino-Catt, Susan J.
 ; APPLICANT: Wang, Hongyu
 ; APPLICANT: Beach, Larry R.
 ; APPLICANT: Wang, Xun
 ; TITLE OF INVENTION: Genes Controlling Phytole Metabolism in
 ; FILE REFERENCE: 0706
 ; CURRENT APPLICATION NUMBER: US/09/118,442B
 ; EARLIER FILING DATE: 1998-07-17
 ; EARLIER APPLICATION NUMBER: 60/055,446
 ; EARLIER FILING DATE: 1997-08-11
 ; EARLIER APPLICATION NUMBER: 60/055,526
 ; EARLIER FILING DATE: 1997-08-08
 ; EARLIER APPLICATION NUMBER: 60/053,944
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 1428
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (118)...(1176)
 US-09-118-442-5

Query Match

Best Local Similarity 0.7%; Score 19; DB 3; Length 1428;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGG 75
 |||||
 Db 6 GTCGACCCACGCGTCCGGG 24

Search completed: July 31, 2003, 23:55:46
 Job time : 188 secs

THIS PAGE BLANK (USPTO)